Ultrafast Time Resolved Spectroscopy Probes Carboxylate Metal Binding and Proton Transfer Kinetics in Condensed Phase

by

Sunayana Mitra

Bachelors in Biochemistry, University of Calcutta, 2008

Masters in Biochemistry, University of Calcutta, 2010

Submitted to the Graduate Faculty of

the Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2021

UNIVERSITY OF PITTSBURGH

KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Sunayana Mitra

It was defended on

July 19, 2021

and approved by

Sean Garrett-Roe, Professor, Department of Chemistry

Sunil Saxena, Professor, Department of Chemistry

Jennifer Laaser, Assistant Professor, Department of Chemistry

Daniel Lambrecht, Assistant Professor, Department of Chemistry and Physics, Florida Gulf Coast University

Ultrafast Time Resolved Spectroscopy Probes Carboxylate Metal Binding and Proton Transfer Kinetics in Condensed Phase

Sunayana Mitra, PhD

University of Pittsburgh, 2021

This dissertation describes two scientific experimental techniques, namely two-dimensional infrared (2D-IR) spectroscopy and time–resolved multiprobe spectroscopy. These methods probe metal coordination geometry in multi-carboxylates and proton transfer kinetics in protic ionic liquids, respectively. Another aspect of this dissertation discusses making the scientific literature accessible to students early in their undergraduate careers.

For the metal–carboxylate structural investigation, symmetric stretch vibrational peaks were probed using the EDTA molecule and divalent Ca^{2+} and Mg^{2+} ions. Multi–carboxylates are components of proteins. The carboxylate symmetric stretching vibration is sensitive to metal–induced structural changes. In EDTA's carboxylate stretching region, the symmetric stretch mixes with the CH bending modes. With 2D-IR spectroscopy, density functional theory, participation coefficient analysis, and energy decomposition analysis calculations, the different vibrational contributions were disentangled and associated with the ion binding geometry. Hence, the ion–binding information from EDTA's symmetric stretch region can be utilized in protein–ion binding research.

In the second investigation, time-resolved multiprobe spectroscopy probed a photoacid dissolved in a protic ionic liquid to unravel the proton transfer kinetics in an ionic liquid environment. Our investigation with this system suggests that proton transport has fast kinetics falling within the Grotthuss mechanism and slow rates that match vehicular transport. The time-resolved multiprobe spectroscopy data, ranging from a few hundred femtoseconds to several nanoseconds, reflect the complicated kinetic pathway of the proton in such a system. This work indicates that protic ionic liquids have complex reaction kinetics that must be modeled using the proton conduction models for water.

Thirdly, this document describes a curriculum that helps undergraduates assess the research presented in scientific literature. The first semester of this curriculum teaches students

to read and comprehend existing original research literature. With worksheet assignments, workshops, and office hours, the students interpret hypotheses, experimental methods, conclusions, and future directions in research. This curriculum provides undergraduates a starting method to evaluate scientific literature and introduces an assignment structure to build on critical thinking skills.

Table of Contents

List of Tables

List of Figures

Dedication

To my first teachers, inspirations, and advocates: Anil Mitra, Subha Mitra, Krishna Sarkar, Late Benoy Sarkar, and Late Sunil Mitra. My strength in completing this journey comes from your blessings.

Preface

As far back as I can remember, I have been fascinated with colors. Hence, I was not surprised to realize my love for lasers one cloudy afternoon in Rochester, NY, in 2012. Thus began my venture to switch my specialization from biophysics to ultrafast spectroscopy. After a year of whirlwind applications to various graduate programs with nonlinear optical spectroscopy groups, it was a pleasure to get an offer from Pitt Chemistry in 2014.

At first, I was apprehensive of my potential as a physical chemist due to my biochemistry background. When I joined Sean Garrett-Roe's laser lab, however, I found my true calling! I thank Sean as I grew to be an independent researcher in his lab, well-versed in nonlinear spectroscopy, with a knack for accuracy, clear communication, and able to pursue my passion. During this journey in Sean's lab, I was introduced to fascinating science, gained new knowledge, met terrific people, and tackled the most challenging problems. Without the thesis committee, progress towards graduation is a difficult hurdle to surpass. Therefore, I thank Sunil Saxena, Daniel Lambrecht, and Jennifer Laaser for their valuable input in my graduation progress.

Next, I thank Eugene Wagner for championing me in my teaching research. One morning at the end of the 2017 Fall semester, I went to his office with my teaching-as-research project idea, and since then, he has supported all my teaching efforts. Dr. Wagner mentored me to grow as a teacher, implement my teaching project, and enjoy facilitating students' understanding of scientific literature. Dr. Wagner's devotion to teaching and drawing lucid connections between chemistry and the everyday occurrence of science is energizing. Under his guidance and unfailing trust, I learned to teach from the undergraduates' perspective and became an efficient educator. Working with him has shaped my teaching and mentoring capabilities.

Paul Donaldson, with whom I have a central laser facility (CLF), RAL collaboration, has a vital role in my research journey. He is a wizard at laser experiments and a marvelous scientist, and I am grateful to have worked with him on several laser experiments. During my first CLF visit, Paul inspired me to keep trying even when things do not go according to plan.

My inspiration to master laser spectroscopy comes from my mentor David W. McCamant. Seeing his group and working on optical projects in his lab, influenced me to learn ultrafast spectroscopy. Dave has cheered me on to perform my best and be courageous during my graduate years.

I am beholden to all the members of the Garrett-Roe group. Being in this group has allowed me to meet immensely supportive, witty, and empathetic colleagues. Samrat, Zhe, Tom, Clinton, Kai (Tycho), CJ, Tyler, and Sydney (Salem & Vlad) thank you for your unflinching support. When I ramble on, your patience with me is exceptional! All the humorous tea breaks, meaningless banter, movie nights, pet adoption talks, crazy jokes during the lunch and dinner outings, late evening chats in the laser lab, and emotional support hold special meaning. Thank you, CJ, for all the refreshing 'Akatsuki time,' conversations and outings! Clinton thanks for patiently explaining cultural references, being understanding, introducing me to all the SciFi series. Tyler and Sydney, special thanks to you for bearing the demanding requests at times and always believing in me! Tyler, thank you for being a good listener. Also, Sydney, you are the 'queen of my heart.' Haoyuan, I am glad that you have joined this lively SGR family.

I am indebted to my seniors Zhe Ren and Thomas Brinzer. From how to operate the equipment and optics to giving feedback in science to eastern culture discussions, and even giving driving lessons! – yes, these seniors did all of that for me! They prepared me for research, taught me everything about the laser lab, and to think about science. I try to emulate their scholarship, mentorship, sincerity, and goodness. Collaborating with Kai Gronborg gave me a listener and someone to bounce off ideas. He can teach the most difficult concepts to any layman with ease. Be the conversation about science, life, food, the latest gizmo, or a comedy series– we can go on for hours! Words are not enough to express my gratitude for the collaboration we have. He has impressed the importance of having fun while programming and keeping one's calm at all times. These colleagues have taught me to conduct science with professionalism and grace during graduate school.

I had the fortune to collaborate with two of the most incredible people: Eric Berquist and Keith Werling. Without them, my understanding of computational chemistry would be dwindling at best! I looked forward to our meetings every week, waiting for the dinosaur, snow

plowing, space-time, and funny blog— talks! We did, however, end up having productive meetings. Whenever I pose some common-scientific-concept questions, their method of teaching at my pace is commendable. Both Eric and Keith have borne all queries with patience, and I am forever grateful for that! Their creativity and curiosity pushed me to be competent in my scientific endeavor.

As for my extended second family in Pittsburgh over the years, I am grateful to all my friends, without whom I could never survive this doctoral adventure! Christina, Rachael, Huiling, Lily, Mina, Soumitra, Haley, Sagnika, Bridget, all my fellow and undergraduate teaching assistants, and students – our conversations, the encouragement, the supportive outlook, and shared experiences made graduate life enjoyable. It gives me great joy to know Christie Hay, our graduate administrator. Her presence and empathy in my graduate life made the journey a lot smoother. Also, I thank Bud Brizuela and Dave Emala for averting many impending computer catastrophes for me! The chemistry department has marvelous people capable of many feats, and I am obliged to so many!

I thank Alekhya for steadfastly being there for me in IISc Bangalore and beyond. Our careers' similar trajectory has eased us to handle our grad school expeditions better together. Thank you to all the UoR friends for sharing your experiences and time when I needed a conversation. I am also grateful to my teacher (Arun Nag) for introducing me to quantum mechanics during my high school days.

My family and kins have my sincere gratitude. Their unconditional love and faith have shaped my zeal for knowledge. My brother (Saayan) and cousin (Samidip) deserve special mention for always being ready for a lively discussion and giving me the courage to move forward in life. Two of my closest relatives, who passed away before they could see me realize this goal, have been with me all along in spirit. Pittsburgh, home to the three rivers, hills, symphonies, art museums, Phipps conservatory, Carnegie libraries, National Aviary, and ever-cheerful people, is a part of me. The nature trails and outings in this vibrant city enabled me to move onward slowly but surely. Finally, I take this opportunity to thank the Universe for getting *'cooler'* everyday and being a place of wonder that always inspires me to be brave, resilient, and dynamic at all times!

1.0 Introduction

Ultrafast vibrational spectroscopy addresses fundamental questions about molecular structure and dynamics. Time–resolved ultrafast techniques interrogate femtosecond to picosecond timescale processes of a chemical or biological system. This thesis will describe two broad categories of time–resolved vibrational spectroscopy, each of which will address two distinct problems in vastly different systems. To elucidate ion–coordination by proteins, we first interrogate ion–carboxylate geometry with two dimensional infrared spectroscopy. Secondly, to better understand the proton transport mechanism in a protic ionic liquid, the kinetics of a proton is monitored with time–resolved multiprobe spectroscopy. A final chapter of this thesis describes a curriculum development, where undergraduates are trained to evaluate primary research.

1.1 Broad Research Goals

1.1.1 Ion–binding

Protein–ion chelation is ubiquitous in all lifeforms. Ion–chelation is critical to protein function. Metals can participate in signaling (*e.g.*, voltage activated Na^+ and K^+ channels), ion–sensing and transport (*e.g.*, Ca^{2+} and Mg^{2+} binding by EF–hand domains), enzymatic reactions (*e.g.*, Fe in cytochromes), and structural functions (*e.g.*, zinc finger proteins).[1–24] Interacting with the ions can change protein structure, enzyme and protein function, or downstream signaling pathways.[25–28] Among these ions, Ca^{2+} is a *second messenger* in G-protein coupled receptor (GPCR) mediated signaling pathways. Almost always, an EF–hand domain resides downstream to the GPCR, which can bind both Ca^{2+} and Mg^{2+} ions. Ca^{2+} and Mg^{2+} ions modulate the structure of the EF–hand domains, which in turn opens or closes these GPCR proteins. Therefore, the structural changes and fluctuations in the EF–hand lead to their ion-sensing properties, global reorganization of structure, and their downstream

structure modulation capability.

Structural changes in such EF–hand proteins can be monitored in real time with two dimensional infrared (2D-IR) spectroscopy. 2D-IR spectroscopy reports the Ångstrom length fluctuations in a system and its environment at an ultrafast timescale–in the range of 100 fs to over a couple of hundreds of picoseconds.[29] The critical factors controlling the system's dynamics are the vibrational lifetime and the molecular environment of the IR chromophore. In this time range, we can observe events such as solvent cage motion, molecular reorientation, ion–induced geometry changes in proteins, and chemical exchange. Utilizing 2D-IR spectroscopy, we look at the earliest local changes in and around the carboxylate reporter group. Correlating that information to the changes in the EF–hand protein can then give insight on their ion–induced structural changes in ultrafast timescales, and how those changes are translated to downstream changes in conformation.

Multi-carboxylates interact with a Ca^{2+} or Mg^{2+} ion in an EF–hand binding pocket. These approaching ions are 'sensed' by the EF–hands. Also, these domains are ion–specific and selective. With ion binding, a structural change starts in the binding pocket, which is translated to an overall alteration of the EF–hand protein conformation. A mechanistic picture of this molecular recognition process, leading to the conformation change of the protein, is yet to be established. Using model compounds that mimic the EF–hand binding pocket is the first step to explain the Ca^{2+} and Mg^{2+} binding interactions. In section 1.2, we will try to understand the EF–hand binding site interactions. Then in chapter 4, we will study EDTA ion–binding to see how it aids in the understanding of these EF–hand multi-carboxylate interactions.

1.1.2 Protic Ionic Liquids

Protic ionic liquids (PILs), room temperature molten salts with a labile proton, are the result of a proton transfer from a Brønsted acid to a Brønsted base. As a subset of a broader class of ionic liquids, PILs have interesting features that separate them from aprotic ionic liquids. The first ever reported and synthesized ionic liquid, ethylammonium nitrate, is a PIL. There are a wide variety of PIL–forming cations and anions, with a range of nanoscale

structure and heterogeneity giving rise to unique properties in the different PILs. Just like aprotic ionic liquids, PILs have polar and non–polar domains that give rise to their molecular heterogeneity. Some properties, like low vapor pressure, thermal stability, ionic conductance, amphiphilic subassembly, and the ability to form hydrogen–bonding networks similar to water, are of great importance in the design of fuel cell replacements.

There is a drive to better understand the proton transfer phenomenon so that we can easily tune a PIL's physicochemical properties for future electrolytic applications. PILs with complete proton transfer show high ionicity and proton conductance, which are necessary processes for a fuel cell application.[30–35] Thus, developing the different PILs for new fuel cell technology requires us to fully understand proton transport in a PIL system. The proton transfer kinetics in PILs is not yet established. Proton transport can occur either by a Grotthuss or by vehicular mechanism. While the Grotthuss pathway involves the hopping of the proton from one molecule donor to the next molecule proton acceptor in picoseconds, in vehicular diffusion the molecule carries the proton in solution at a much slower nanosecond proton conduction rate. Not understanding the exact nature of proton conduction in PILs hinders the implementation of these solvents as future battery resources. Therefore, in section 1.3, we will learn more about the proton transport phenomenon, and in chapter 5, we will explore the proton transport in a protic ionic liquid system using an initiator photoacid.

1.1.3 Curriculum Development

Undergraduates will benefit if they can evaluate scientific research at an early stage of their degree. Comprehension of the different research areas will help undergraduates utilize techniques and make advances in applications. Therefore, undergraduates must be able access these concepts. This leads us to the last section, where we will outline a teaching curriculum that was developed during the course of my research. This curriculum prepares them to read and evaluate primary research and extrapolate. Section 1.4 will introduce the need for academic institutions to better equip undergraduate students to understand scientific literature. Utilizing the variety of experimental techniques presented in the different research publications will help undergraduates to build their own critical thinking repertoire. So, in

chapter 6, we will present a curriculum aimed to develop these skills in an undergraduate student.

1.2 Multi-carboxylate Ion Binding

2D-IR spectroscopy informs us about the bond–length and picosecond resolution events. The binding loop in EF–hands undergo ultrafast fluctuations which contribute to the ion–induced structural changes in the protein. These fluctuations are sensitive to the overall changes in the environment such as an approaching ion. The ultrafast time resolution of 2D-IR spectroscopy provides a way to investigate those changes. Studying those picosecond fluctuations with 2D-IR spectroscopy offers better understanding of the ion–sensing by these EF–hands than NMR and EPR spectroscopies or stopped flow techniques. EPR spectroscopy cannot look at changes faster than µs and require invasive spin labels. NMR techniques have been used to study EF–hands but cannot probe faster than ms timescale required for the binding loop examination. Single molecule techniques also have been used to study these EF–hands but fail to offer the picosecond time resolution and also require a large non-native dye insertion, thereby disrupting the structure. Stopped flow methods can measure rapidly up to 100 µs and are incapable of inspecting the binding loop dynamics. With 2D-IR spectroscopy, the information in these binding loop multicarboxylates can be unpacked, and we can extract both structural changes and ultrafast conformational flexibility of these proteins.[36–47]

This section will now expand on the importance of understanding the ion–induced structural changes in the EF–hand multi-carboxylate domains. The broad goals above mention that Ca^{2+} and Mg^{2+} ion–binding triggers the structural changes in EF–hand, which regulates functional responses in the downstream signaling scheme. EF–hand–Ca^{2+}–binding proteins always contain more than one EF–hand domain. An EF–hand motif has a helix–loop–helix structure, with 12 residues per helix and about 29 residues per domain. The canonical EF–hand binding loop starts and ends with carboxylate side–chain amino acid residues, aspartate and glutamate, respectively.[48]

X-ray crystal structures of several EF–hand–containing proteins show that these domains

function in pairs. A hydrophobic residue (usually isoleucine, valine, or leucine) in the secondary structure of the ion–binding loop hydrogen bonds to the other EF–hand, maintaining the structural integrity of the pair. An exclusive EF–hand domain unpaired to another is rarely seen in nature and almost always has diminished Ca^{2+} binding capability. This suggests that cooperativity between the domain pair plays a role in ion–selectivity, affinity, and specificity.[49]

The effect of ion–selectivity can be better understood in the context of the cellular environment. Intra-cellular Ca^{2+} is in the $\sim$ 100 nM range, while intra-cellular Mg^{2+} is in the $\sim$ 2 mM range. Therefore, the EF–hand proteins encounter an intra-cellular $Ca^{2+} : Mg^{2+} = 1 : 1000$ concentration ratio, whereby the Mg^{2+} ion concentration dominates over Ca^{2+} in the cytoplasm. Despite this significant concentration disproportionality between the two ions, the EF–hand proteins are highly selective to Ca^{2+}.[50] To explain this selectivity and specificity, a thorough understanding of the binding loop structure is necessary.

The composition and length of the Ca^{2+} binding loop in EF–hand proteins can widely vary, from being 12 residues in parvalmumin to a 20 residue binding loop in the neuronal calcium sensor proteins. In all these variants of the EF–hand loop, it is the N–terminal domain amino acid residues that get replaced, whereas the C–terminus composition remains constant in length. Despite the variety in the loop length and composition, the pentagonal bipyramidal binding geometry, the Ca^{2+} binding affinity, and the Ca^{2+} and Mg^{2+} specificity is preserved. This pentagonal bipyramidal geometry of the loop is a byproduct of one backbone carbonyl coordination, the four side chain oxygen and one carboxylate interactions, and one ordered water molecule ligand. Additionally, the last glutamate residue in the binding loop, that donates a carboxylate side–chain, is conserved in all EF–hand structures.[51]

The multi-carboxylate interaction in the EF–hand binding loop utilizes seven oxygen atoms to bind a Ca^{2+} ion in the pentagonal bipyramidal Ca^{2+} binding geometry. As mentioned earlier, the N-terminus of the Ca^{2+} binding loop can vary, having changeable amino acid residues and length. The N–terminus does not disturb the bipyramidal geometry. This is not true for the C–terminus, where a residue change almost always disrupts the pentagonal bipyramidal Ca^{2+} coordination. These carbonyl oxygens bind the Ca^{2+} ion in a *unidentate* coordination, and the glutamate in the last position binds in a *bidentate* geometry.[52]

Infrared spectroscopy is extensively used to determine the different EF–hand ion–binding

geometries. Existing literature has utilized the carboxylate stretching vibration to understand the unidentate, bidentate, bridging, and pseudo–bridging coordination for EF–hand proteins. The carboxylate antisymmetric stretch, with a high absorption coefficient, is the preferred mode for studying the structure of these proteins. However, the peakshape changes in the antisymmetric stretch upon Ca^{2+} and Mg^{2+} coordination is small. Oftentimes, the energy gap between the symmetric and antisymmetric stretch is also used to elucidate EF–hand ion–coordination structure.[53,54] This correlation between the spectrum and the ion–binding geometries was developed in the acetato-carboxylate complexes in the mid–1950s, and all these empirical structure relations are for single carboxylate coordination. In chapter 4, we will detail how to utilize the complex symmetric stretch region of multi-carboxylates to elucidate ion–binding conformations.

After an ion binds to an EF–hand motif, the changes in the infrared peaks are more evident in the carboxylate symmetric stretch than the carboxylate antisymmetric stretch. In all the previous work, the origin of the vibrational structure in the EF–hand protein's carboxylate symmetric stretch region remained unresolved. Mostly the carboxylate antisymmetric stretch and the energy gap between the two stretches are used to decipher the ion–binding geometry. After a Ca^{2+} or Mg^{2+} ion binds the peak structures in the symmetric stretch region, clearly reflect the specific ion–induced geometry changes in the new peaks. Thus, to comprehensively explain these binding–induced conformational changes, we must understand the vibrational shapes of the carboxylate symmetric stretch region.

To elucidate the protein–ion binding pocket interaction in an EF–hand domain, we can use model compounds. Multi-carboxylate model compounds can mimic the ion interactions seen in the EF–hand binding pocket. By first analyzing how the binding pocket interacts with an ion, we can deduce how that structural change is relayed to the rest of the EF–hand domain. Such a focused scrutiny on the active site will reflect the interplay between the carboxylate residues and the ion in the pocket, providing insights on ion–sensitivity and selectivity. Ethylenediaminetetraacetic acid (EDTA) is such a model compound, which binds divalent ions with its tetra–carboxylate groups. The heptadentate coordination of the EF–hand binding loop is closely modeled by the hexadentate chelation of EDTA. The carboxylate symmetric stretch region of EDTA is rich in structure, in the presence or absence

of ions. In chapter 4, the origins of these vibrational structures will be correlated to the binding geometry of Ca^{2+} and Mg^{2+} ions.

Understanding the binding induced changes in real time is essential for elucidating properties such as ion–sensing, ion–specificity, ion–selectivity, and global structural reorganization in these EF–hand proteins. X–ray crystallography provides high resolution static structures of these Ca^{2+}– and Mg^{2+}–binding proteins. NMR[55–57] and EPR[58] spectroscopies are unable to give insights on the fast timescale structural changes of these proteins, while stopped flow methods are unable to resolve the ultrafast timescale changes in these EF–hand proteins.[59–62] 2D-IR spectroscopy can interrogate both dynamics and structure of these multi-carboxylate ion–binding domains.

2D-IR spectroscopy can reveal protein structure and dynamics with cross-peaks, bond length specific spatial resolution, and sub–picosecond temporal resolution. The 2D-IR chromophores can be either internal or external. Internal probes, such as Amide I–IV backbone vibrations, report the global structures of proteins.[40–47] External probes, such as nitriles[63], azide[64], thiocyanate[65], or isotope labels in CH, COO, and CN residues of proteins[66] also elucidate structure and dynamics. These external probes can be placed inside a protein sequence either by chemical alteration or by inserting artificial amino acids in regions of interest.[67–70] Placing an unnatural amino acid residue inside a protein has also been utilized to determine structural and dynamical information.[71–73] Couplings, lineshapes, and dynamics of a K^+ voltage gated channel residing in the membrane reflect the structural changes in the protein upon voltage gating.[74]

Combining 2D-IR spectroscopy with molecular modeling, MD simulations, and density functional theory (DFT) calculations yields atomistic insights into protein structure and dynamics. Using sophisticated spectroscopic maps, the 2D-IR spectra of folded, disordered, globular, fibrillar, and multisubunit protein structures and dynamics can be modeled and replicated, correlating the vibrational structure to the protein function. Simulating the couplings between different vibrations, from model Hamiltonians, has aided in the interpretation of many complex protein 2D-IR spectra. Since the past decade, MD simulations have had a critical role in reproducing the 2D-IR spectra from the individual trajectories, leading to better understanding of the protein dynamics.[75–91]

Utilizing side–chain carboxylates as internal 2D-IR chromophores informs us about the localized binding–induced changes. Local binding changes require probes that are site specific, such as amino acid side chains in the region of interest. Side chain groups are sensitive to specific events in localized positions, such as metal binding, hydrogen bonding environment, or salt bridges in proteins.[92,93] 2D-IR spectroscopy of the carboxylate antisymmetric stretch, which is a local vibration, reports EF–hand–metal active site structural distortions and mutation effects.[94,95] Therefore, 2D-IR spectroscopy can probe these ion–induced active site coordination changes, opening up the possibility to examine complicated EF–hand mediated downstream structural changes. Chapter 4 will detail how couplings from the cross-peaks in a multi-carboxylate 2D-IR spectrum tells us about the different ion–binding geometries. This structural knowledge will facilitate our understanding of how an EF–hand binding to a Ca^{2+} or Mg^{2+} triggers a chain of downstream functional responses.

In chapter 4, we will learn how 2D-IR spectroscopy combined with participation coefficient analysis can furnish binding geometry information. The carboxylate symmetric stretch region is free of overlap from the amide vibrations common in proteins, thereby offering a structurally sensitive region with the least amount of spectral congestion to interrogate structural changes. Focusing on that carboxylate symmetric stretch of EDTA, we show that this region of the vibrational spectrum is rich in structural information due to the overlap with CH bending vibrations. We then show how to unscramble that information by developing a participation coefficient methodology. Finally, we will assign the peaks in the linear IR spectra, determine the origin of the couplings in the 2D-IR spectrum, and correlate the spectral changes to metal–coordination induced changes in EDTA. This EDTA research opens a new part of the spectral window that can be used as a 2D-IR chromophore that was unavailable before. Moreover, the uncongested carboxylate symmetric stretch region, the participation coefficient analysis method along with the ultrafast time resolution of 2D-IR spectroscopy yields a region of the spectrum that can cleanly probe and is sensitive to the binding loop fluctuation and ion–sensing by EF–hand proteins. In the future, this comprehensive spectroscopic tool developed through this research will allow us to explore binding geometry structure and dynamics of more complex EF–hand proteins.

The research conducted in chapter 4 adds the carboxylate symmetric stretch region as a

probe to the wide variety of vibrational bands available for 2D-IR spectroscopy of protein structure and dynamics elucidation. Chapter 4 will reveal the origin of vibrational bands in the carboxylate symmetric stretch of condensed phase EDTA in the presence and absence of ions. New peaks and peak shape changes in the carboxylate symmetric stretch indicate the structural changes in EDTA as metal binds. This research opens the possibility to explore ion–binding in more complicated biophysical systems, such as Ca^{2+} and Mg^{2+} binding EF–hands utilizing 2D-IR spectroscopy and electronic structure theory of the carboxylate symmetric stretch region.

1.3 Proton Kinetics in Protic Ionic Liquids

This section unravels the different methods contributing to the understanding of proton transport in PILs. We will also see how techniques applied to understand water's proton transfer can help us examine the PIL systems. Proton transfer from the acid to base determines the hydrogen bonding ability of PILs.[96,97] With a strong acid and base combination, each with their respective pK_a values, there is an increase in the proton transfer capacity of the PIL. The pK_a values come from the constituent acids and conjugate acid of the respective base dissolved in water. Thermodynamics of proton transfer can be approximated from these pK_a values. In some liquid PIL, a high aqueous ΔpK_a value results in a complete proton transfer from the acid to the base, whereas in a solid state PIL, a low ΔpK_a value can complete the proton transfer process. Therefore, ΔpK_as obtained from the aqueous medium acid and base starting materials are inaccurate representation of the proton transfer process. ΔpK_a values do not fully reflect the ionic contributions due to the structural heterogeneity of the cations and anions in PILs.[98–100]

The nature of the cation, the chain length, and the functional groups in both cations and anions all affect the proton transfer process.[101,102] Cations that are commonly used in PIL synthesis are primary, secondary, and tertiary ammonium ions, mono– and di–imidazolium ions, and guanidium ions, while nitrate, formate, acetate, hydrogen sulfate, and trifluoroacetate are frequently used anions. A Walden plot can approximate the proton

transfer in newly synthesized PILs. An *ideal* line in a Walden plot is created from a reference KCl salt that undergoes complete ionization. The closer the acid dissociation constant (pK_a) is to the *ideal* line in a Walden plot, the more complete is the proton transfer in the PIL. Primary amines are closer to this line than secondary and tertiary amines, demonstrating a near complete proton transfer process in the former.[103] Using the Walden plot, PILs with different cations and anions have been investigated for their proton transfer characteristics.[96,97,101,104] Walden plots utilize the acid dissociation constants from aqueous solutions to characterize the ionicity in PILs, however, they fail to characterize if a low ionization occurs from ion pairing or a deficit in proton transfer.[103] Even though aqueous ΔpK_a and Walden line depictions are ambiguous metrics of the proton transfer process, they are conveniently used to determine the conduction mechanism of the proton in PILs.

It is yet to be determined if PILs have a Grotthuss or vehicular mode of proton transport or a combination of both these mechanisms. The fast proton transfer, in a Grotthuss hopping process, induces a rearrangement in the hydrogen bond network. A study of the far-IR region, which contains the low frequency intermolecular stretching and bending of hydrogen bond modes, revealed the similarities in hydrogen bonding structure of both water– and ammonium– based PILs. This finding along with DFT calculations indicate the strong resemblance between the hydrogen bond network of water and PILs.[105] Johnson *et al.*[106] have further established that ethylammonium nitrate and water have comparable dynamical heterogeneity and fluctuations in their hydrogen bond network.

Proton transfer reactions require a free H^+ in solution. When the proton is a component of the donor or acceptor molecule, it might be difficult to track its motion initially. To circumvent this problem in a PIL solution, we can introduce and track an external proton from a photoacid molecule. Photoacids can initiate the excited state proton transfer process.[107,108] Photoacids have low acid dissociation constant when in the excited state. This low pK_a of the photoacid ensures that it is a strong acid in the excited state, which can release a proton into solution, whose movement is monitored with spectroscopic measurements. In practice, the photoacids used in proton transfer studies have extensive aromatic ring structures. Initial absorption of a photon by the photoacid triggers a charge transfer to its ring system, leading to a charge redistribution in the molecule. This redistribution in the excited state decreases

the pK_a and a proton is released. Solvent influences a lot of the photoacid characteristics by modulating the charge reorganization after photoexcitation. Therefore, isotope content, presence of salts, water concentration, solvent polarity, and temperature regulate the pK_a of the photoacid and its excited states in any particular solvent.

Water, which has a similar hydrogen network as PIL, can facilitate our understanding of proton transfer in PILs. Acid catalysed proton conduction in water is a picosecond timescale event and is extensively studied, as proton transfer reactions in water play significant roles in chemical reactions and biological pathways.[109–115] In water, conduction of the proton occurs via two pathways: vehicular and Grotthuss. In the diffusion (vehicular) pathway, the molecule moves with the proton in solution in slow nanosecond timescales, while in the Grotthuss pathway the proton hops from one molecule donor to the next molecule proton acceptor in picoseconds. The H_3O^+ (Eigen cation), $H_5O_2{}^+$ (Zundel cation), and $H_9O_4{}^+$ (highly solvated Eigen cation) participate in the Grotthuss type proton hopping motion, which results in the hydrogen bond network rearrangement of water.[116] In the Grotthuss pathway, where the mobile entity is a shift in the hydrogen bond partner, the proton charge can travel great distances in solution via the extensive hydrogen bond network.

Utilizing time-resolved spectroscopy techniques, Erik Nibbering's group has pioneered in understanding the complex kinetic mechanism by which the released proton journeys through an aqueous acidic solution.[117] Doing so has revealed the different categories of intermediate complexes of varying sizes, depending on the number of solvent molecules attached to them. Mohammed *et al.*[118,119] investigated the effect of proton transfer in dilute and concentrated quantities of acetate aqueous solution and found that several intermediate complexes participate in the Grotthuss type proton conduction. Proton transfer in water proceeds in a sequential manner, with the proton hopping from an acid donor to a base acceptor via intermediate water bridges. These water bridges are structurally diverse and can be just one hydronium to two hydronium ions or can extend to a chain of H_2O molecules carrying the positive charge. These femtosecond time-resolved experiments have revealed the existence of 'tight–complexes', 'loose–complexes', and 'solvent switches' in aqueous solution.

Ultrafast transient mid-IR spectroscopy is largely used to follow the proton trajectory in water. Both Grotthuss and vehicular proton transport is found to be prevalent in aqueous

solutions. In water, the complete kinetics of acid–base neutralization models consider the participation from all the donors, acceptors, and the solvent molecules in the relevant shells. There can be a variety of loose complexes and solvent switches in solution, each with a different number of connecting water molecules. These bridges of water molecules can be thought of as 'proton wires' conducting the charge through solution. Each of these loose complexes and solvent switches have their distinct donor and acceptor kinetics.

In an acidic aqueous solution, whether the proton transfer mechanism will be Grotthuss or vehicular depends on the type of intermediate complexes formed.[117] In all these species, the criteria for the proton to be passed on from the donor to acceptor depends on the absolute or relative reactivities, the geometry of the reactants, the number of water molecules between the acid and the conjugate base as well as their relative orientation. Proton transfer in tight complexes, where the donor and acceptor are contact pairs, follows a Grotthuss mechanism with minimum disruption of the solvent shell around the reactants. In loose complexes, the solvent shells have a major contribution in the proton transfer process. As the acid donor and the base acceptor is separated by a string of water molecules, the proton can traverse a large distance through the solvent and be stabilized by the water molecule bridge between them, or the solvent shell between the acid and base can reorganize to create a tight complex and ultimately transfer the proton at ultrafast rates. When the concentration of acceptor is low, the proton in the system will diffuse through the solution to reach a base molecule and diffusion limited kinetics is observed.

Following the footsteps of the Nibbering group's work on proton kinetics in water and implementing the technology on PILs will give a comprehensive picture of the proton transfer mechanism in PILs. To inspect the proton transport mechanism in PILs, we will employ time–resolved multiprobe spectroscopy (TR^MPS)[99] in chapter 5. With TR^MPS, we can observe the proton kinetics from 100 fs to several milliseconds.[120] TR^MPS can track both the ultrafast Grotthuss hopping of the proton and report slower nanosecond vehicular proton transfer process. Chapter 5 will detail how TR^MPS uses 400 nm UV light to initiate the proton release from the photoacid into solution and then probes the PIL solution at different timescales to track the H^+ trajectory. We interrogate the proton transfer kinetics in ethylammonium formate by photo–ejecting a proton into solution using the photoacid pyranine. A careful

evaluation of the vibrational bands of the photoacid and the acceptor anion from the PIL aids in the understanding of the proton conduction process. The analysis of this transient experiment gives us insight on the kinetics of the proton transfer mechanism in PILs.

1.4 Undergraduate Curriculum Development

In this section, we will introduce the need for a curriculum that was developed as as part of this thesis. Gaining an undergraduate degree in any academic institution implies that a student is developing their analysis, evaluation and synthesis skills. This educational process is supposed to help students develop the ability to review and analyze information and draw meaningful conclusions. It is a long term endeavor that requires regular practice in order to become proficient. Hence, students need to start understanding the process of scientific research early. Then they can evaluate primary research efficiently by the end of their undergraduate career. Survey data from students in our honors general chemistry course indicate that students are quite aware that research documented in the literature significantly contributes to the development of the different existing technologies. They even understand the importance of research literature in shaping their education, future career, and understanding of the world. Skillful reading of research journals requires students to master the art of understanding the terminologies and concepts presented in the primary literature. It takes much practice to gain expertise in the efficient evaluation of academic literature. Therefore, an early start to familiarize undergraduates to original research literature is imperative.

Students need to grasp the differences between original research and secondary literature in the process of getting acquainted with scientific literature. Both primary and secondary articles have a specific order of documenting research, and the students need to be aware of the nuances. For instance, broad goals and broad conclusions are found in both secondary and primary article, whereas the specific hypotheses and results are exclusive to primary research papers. Besides, the students should gain the ability to connect their textbook concepts with cutting-edge research in real-time. From their freshman year, the students should recognize

the interconnections between the various aspects of their course work concepts and original research. This mental exercise will hone their critical thinking skills. An undergraduate curriculum is comprehensive only when a student can successfully apply the knowledge from a course and extrapolate it to generate new ideas. In chapter 6, we will show a new scaffolded approach to introduce undergraduates to scientific literature early in their first year. This method is tested on the honors general chemistry I course and shows promising results on the student research performance in class and later in the semester.

2.0 Theory

Theory in this chapter will aid in the understanding of concepts found in chapter 4. The first two sections, 2.1 and 2.2, in this chapter are based on the textbook *Concepts and Methods of 2D Infrared Spectroscopy* by Hamm and Zanni.[29]. Concepts in these two sections will build our understanding of the origin of a two-dimensional infrared (2D-IR) spectrum from ultrafast pulses and the structural information we gather from cross-peaks in a spectrum. In section 2.3, density functional theory is discussed. This will help us comprehend some of the computational calculations used in this thesis.

2.1 Two-Dimensional Infrared Spectroscopy

2.1.1 System Perturbation with Radiation

Spectroscopy aims to discern the interplay between matter and electromagnetic radiation. Picosecond to femtosecond interactions between the molecule and its environment are probed by nonlinear ultrafast spectroscopic methods. 2D-IR spectroscopy, a third order nonlinear technique, utilizes three electric field interactions with the molecule to interrogate its structure and dynamics.[29] In 2D-IR spectroscopy, a semi–classical approach is commonly used to interpret this interaction where the electric field is treated classically and the vibrational states of the system are considered to be quantum mechanical in nature. The classical electric field, $\vec{\varepsilon_t}$, is

$$\vec{\varepsilon(t)} = \varepsilon'(t)\cos\left(\vec{k}\cdot\vec{r} - \omega t + \phi\right), \tag{1a}$$

$$2\vec{\varepsilon(t)} = \varepsilon'(t)(e^{-i\omega t} + e^{+i\omega t}) = \varepsilon(t) + \varepsilon^*(t), \tag{1b}$$

where in equation 1a, $\varepsilon'(t)$ is the pulse envelope, $\vec{k}$ is the wavevector with information about direction of light propagation, ω is the frequency, and ϕ is the phase. $\vec{\varepsilon_t}$ is then represented as a sum of the real and the complex components of the electric field (equation 1b).

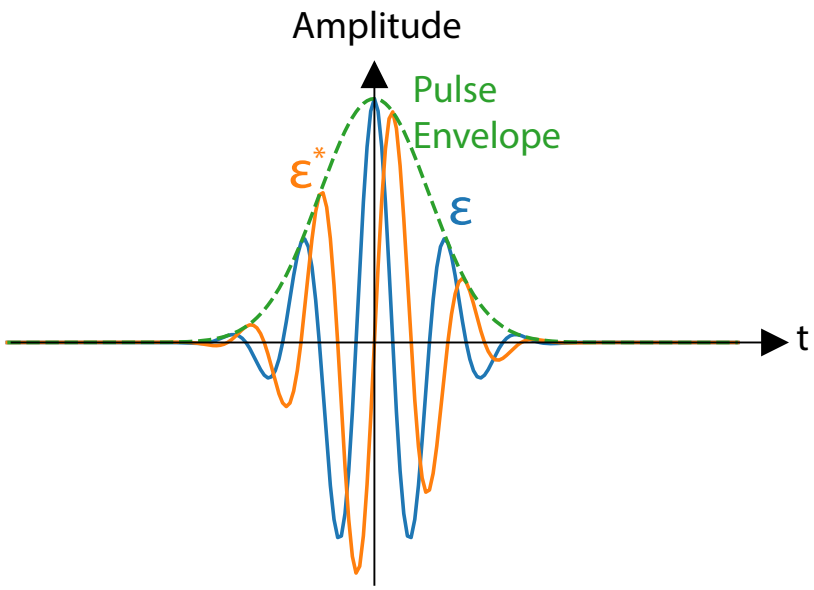

Figure 2.1: The pulse envelope includes both the electric field ε and its complex conjugate ε^*.

With a time-dependent wavefunction, $\psi(t)$, and Hamiltonian, $\hat{H}(t)$, for our system, we can describe its evolution from the instant of perturbation using the time-dependent Schrödinger equation.[121,122] For 2D-IR spectroscopy the interaction between the radiation and the system is weak; *i.e.*, the laser pulse excitation is a weak driving field, so it is sufficient to consider the time-dependence after a perturbation. Then we write the total Hamiltonian, $\hat{H}(t)$, as a combination of the time-independent Hamiltonian and the time-dependent perturbation.

$$\hat{H}(t) = \hat{H}_0 + \hat{W}(t), \tag{2}$$

where

$$\hat{W}(t) = -\hat{\mu}\varepsilon(t). \tag{3}$$

Here, the external laser pulse field is represented by the scalar valued function $\varepsilon(t)$ and $\hat{\mu}$ is the molecular transition dipole moment. $\hat{\mu}$ is the expectation value of transition between two unlike states, $m \to n$. By calculating the relevant off-diagonal elements of the dipole

matrix, which is a sum of the product of all the charges and the electron positions, we get the transition dipole integral,

$$\hat{\mu}_{nm} = \int \psi_n^* \hat{\mu} \psi_m = \langle m| \hat{\mu} |n\rangle . \tag{4}$$

As the laser pulse strikes our molecular system, the latter goes into a non–equilibrium state. From the moment of system perturbation by a pulse, we measure the evolution of time-dependent coefficients, which eventually give the energy gap between the initial (E_m) and final (E_n) states. Vibrational spectroscopists tend to look at energy differences in a system in the units of frequency, so it is convenient to define the $\omega_{nm} \equiv \frac{E_n - E_m}{\hbar}$. After substituting the time-dependent perturbation with equation 3, we can recover the effect of the *transition dipole moment* on the time-dependent coefficients and gain information on the frequency, ω_{nm}. The transition dipole moment (TDM, equation 4) is a combination of transition dipole strength, $\frac{d\mu}{dx}$, and the coordinate of the vibrating bond, $\hat{x}$,

$$\langle m| \hat{\mu} |n\rangle = \frac{d\mu}{dx} \langle m| \hat{x} |n\rangle . \tag{5}$$

The latter, $\langle m| \hat{x} |n\rangle$, is responsible for the vibrational selection rules, where $n = m \pm 1$. Throughout this text, the description of the TDM is considered time-independent; however, in some situations where the orientation of the TDM affects the system, its time dependence becomes important in understanding the system's evolution.[122–124]

When a single wavefunction can describe the system, it is called a *pure state*. Disorder in an ensemble does not allow us to describe a system with one wavefunction. To fully characterize a condensed phase system, an ensemble, we need the density matrix formalism $\hat{\rho}(t)$. In the ensemble density matrix representation, the diagonal elements are the *population states*. In a population state, the ensemble average resides in one of the energy levels of the system. The off-diagonal elements of the density matrix are the *coherence states*, which give rise to the emitted field in the observed system. In the coherence states, the molecular ensemble is in a time-dependent linear superposition state between two energy levels inside the density matrix. In a sense, the density matrix defines the eigenstates of the molecular

ensemble as a linear combination of the initial pure states. The density matrix evolves with time according to the Liouville-von Neumann equation,

$$i\hbar\frac{d}{dt}\hat{\rho} = [\hat{H}, \hat{\rho}], \tag{6}$$

where the $[\hat{H}, \hat{\rho}]$ is the commutator of the Hamiltonian and the density matrix operators.

We can expand the density matrix in the interaction picture using the perturbative power expansion and obtain its time evolution, which gives a nested commutator expression for the density matrix. The nested commutator contains information about the evolution of the molecular ensemble after the perturbation and thoroughly lays out the effect of the transition dipole from the side of the electric field and its complex conjugate. Each of these interactions are easier to track when using *double sided Feynman diagrams*, which organize the interactions, time-intervals, states of the system, and absolute times.[121–124]

2.1.2 Macroscopic Polarization and Molecular Response

In the non-equilibrium state created by the laser pulse interaction, the molecular ensemble eigenstates become a linear combination of the initial states. After the laser pulse, when the molecular ensemble is in a linear superposition state, the time dependence of that state reflects the time dependence of the unperturbed system's Hamiltonian. This is termed as the *molecular response*, $R(t)$. The system's charges oscillate after the laser pulses, to produce the $R(t)$. The *macroscopic polarization*, $P(t)$, is the convolution of the laser pulse electric field with the molecular response $R(t)$. This molecular response represents the ensemble average molecular response of the whole system and Liouville-von Neumann equations can characterize such $R(t)$ and $P(t)$ functions. Experimentally we measure the emitted signal field, which depends on $P(t)$.

$P(t)$ comes into play due to the presence of a random distribution of molecules in the system, which are vibrating in phase after a laser pulse interaction. $P^{(n)}(t)$ is the expectation value of the dipole operator, $\hat{\mu}$ (equation 7). In a nonlinear electric field interaction, the n^{th} order expression for macroscopic polarization is the trace of the n^{th} order density matrix acting on the TDM operator. Expanding that expression we first get the convolution of the n

electric fields in absolute time units, τ_n. Secondly, the $P^{(n)}(t)$ contains a nested commutation between the density matrix and the transition dipole operator, which is a byproduct of the Liouville-von Neumann treatment of the ensemble density matrix.

$$\begin{aligned} P^{(n)}(t) &= \langle \hat{\rho}^n(t)\hat{\mu}\rangle \\ &= -\left(\frac{-i}{\hbar}\right)^n \int_{t_0}^{t} d\tau_n \int_{t_0}^{\tau_n} d\tau_{n-1} ... \int_{t_0}^{\tau_2} \varepsilon_{\tau_n}\varepsilon_{\tau_{n-1}} .... \varepsilon_{\tau_1}\, d\tau_1 \\ &\cdot \langle \mu(t)[\hat{\mu}_I(\tau_n), [\hat{\mu}_I(\tau_{n-1}), ...[\hat{\mu}_I(\tau_1), \rho(t_0)]...]]\rangle. \end{aligned} \tag{7}$$

After changing the absolute times (τ_n) to time intervals (t_n) and accounting for the $\rho(t_0)$ initial Hamiltonian, we convert the $P^{(n)}(t)$ expression with a nested commutator of the transition dipole interactions and the convolution of n electric fields into units of relative time, t_n. Transforming this nested commutator into the n^{th} order nonlinear response function, we get the expression

$$\begin{aligned} P^{(n)}(t) = \int_0^{\infty} dt_n \int_0^{\infty} dt_{n-1} ... \int_0^{\infty} dt_1 \varepsilon(t-t_n)\varepsilon(t-t_n-t_{n-1}) .... \varepsilon(t-t_n-...-t_1) \\ \cdot R^{(n)}(t_n, t_{n-1}, ..., t_1). \end{aligned} \tag{8}$$

Embedded inside equation 8 is the convolution expression of the n^{th} order nonlinear response function. Here, $R^{(n)}$ represents the nonlinear response from all the molecules in the ensemble that the density matrix describes. The complete $R^{(n)}$ is given by,

$$R^{(n)}(t_n, t_{n-1}, ..., t_1) = -\left(\frac{-i}{\hbar}\right)^n \langle \mu(t_n + t_{n-1} + ... + t_1)[\hat{\mu}_I(t_{n-1} + ... + t_1), ...[\hat{\mu}(0), \rho(-\infty)]...]\rangle, \tag{9}$$

where the $\rho(-\infty)$ is the ground state density matrix before any laser pulse interaction.

$P^{(3)}$ is a convolution of the three nonlinear electric fields with the third-order response. As the laser pulse electric field pushes and pulls on the molecular charges, it makes the molecule vibrate, giving rise to the system specific molecular response $R^{(3)}$. The phase of the laser pulse and that of the resonantly excited vibrations in the ensemble are synchronized with each other. Macroscopic polarization ($P^{(n)}$) is the non-equilibrium charge distribution in the system due to the molecules interacting with the n laser pulses. The evolution of the macroscopic polarization over time reveals structural and dynamical information about the molecule. In 2D-IR spectroscopy, $P^{(3)}$ provides information about the molecular structure,

the energy exchange between the molecules, the inter-molecule and solvent-molecule couplings, and fluctuations in the solvent environment.

The third-order nonlinear response, $R^{(3)}$, that a 2D-IR experiment generates can be expressed as

$$\begin{aligned} R^{(3)}(t_1,t_2,t_3) = -\left(\frac{-i}{\hbar}\right)^3 Tr(\hat{\mu}_3\hat{\mu}_2\hat{\mu}_1\hat{\mu}_0\hat{\rho} - \hat{\mu}_3\hat{\mu}_2\hat{\mu}_1\hat{\rho}\hat{\mu}_0 - \hat{\mu}_3\hat{\mu}_2\hat{\mu}_0\hat{\rho}\hat{\mu}_1 \\ +\hat{\mu}_3\hat{\mu}_2\hat{\rho}\hat{\mu}_0\hat{\mu}_1 - \hat{\mu}_3\hat{\mu}_1\hat{\mu}_0\hat{\rho}\hat{\mu}_2 + \hat{\mu}_3\hat{\mu}_1\hat{\rho}\hat{\mu}_0\hat{\mu}_2 + \hat{\mu}_3\hat{\mu}_0\hat{\rho}\hat{\mu}_1\hat{\mu}_2 - \hat{\mu}_3\hat{\rho}\hat{\mu}_0\hat{\mu}_1\hat{\mu}_2), \end{aligned} \tag{10}$$

where $\hat{\rho} \equiv \hat{\rho}(-\infty)$ and $\hat{\mu}_n \equiv \hat{\mu}(\sum_{j=1}^{n} t_j)$. Equation 10 has eight terms in its nested commutator. When explicitly written out, equation 10 includes interactions of both the electric field ε and its complex conjugate ε^* giving a total of 8^2 terms. This ordering scheme is visualized by the Feynman diagrams described in Hamm and Zanni.[123] Out of these 64 terms, for a single oscillator system that is initially in the ground state, only 12 pathways survive after applying the *Rotating Wave Approximation*[121,124] on them. To describe a single vibrational mode, we use only 6 of them (three *rephasing* and remaining three *non-rephasing*), as six of the Liouville–von Neumann pathways are the complex conjugates of each other.[122,123]

When one uses ultrafast femtosecond laser pulses (∼50-150 fs) for vibrational spectroscopy measurements, the contribution of the laser pulse envelope to the third-order nonlinear polarization becomes negligible compared to the molecular response and we are within the *semi-impulsive limit* (Figure 2.1). Here, when the molecular timescales are slower than the laser pulse, one assumes that the laser pulse is a δ–function compared to any timescale relevant to the dynamics of the molecules and longer than the oscillation period of the laser light. A laser pulse within the semi–impulsive limit will yield an n^{th}-order molecular response that equals its macroscopic polarization. 2D-IR spectroscopy utilizes these ultrashort femtosecond laser pulses that are smaller than the timescales significant to the molecules. Therefore we measure the emitted signal with the correct phase, wavevector, and carrier frequency coming from the $P^{(3)}$ term. Within this δ–function approximation, the emitted signal becomes directly proportional to the third order nonlinear response, $R^{(3)}$.

2.1.3 2D-IR Signal

An emitted signal needs to be measured along with all the phase dependencies of the $R^{(3)}$. A *local oscillator* can be used to detect the emitted signal. For example, the *box-CARS* geometry uses a fourth pulse as the *local oscillator* (LO). In the pump-probe geometry, the third pulse assumes the role of the local oscillator. The probe pulse both creates the signal field and *heterodynes* with it, hence the emitted signal is called *self-heterodyned*. In the pump-probe geometry, the probe locks the time and phase information of the intrinsic local oscillator, which is used to recover the same information about the response functions.

Terms such as *homodyne* and *heterodyne* detection are often used in ultrafast vibrational spectroscopy. In *homodyne* detection, a square–law detector measures the intensity of the signal field, $I = |E_{sig}^2|$. As the signal intensity gets averaged over time, the *homodyning* leads to a loss of phase information. *Heterodyning* an emitted field means that another ultrashort pulse along with the signal field is incident on the square law detector. This creates an interference term between the signal and the extraneous ultrashort pulse, $S \propto |E_{LO} + E_{sig}|^2$, which falls on the detector.

By heterodyning the emitted signal we get the time and phase dependence of the signal field, which helps in the Fourier transform and generation of a 2D-IR spectrum. The total signal, S, falling in the detector includes three terms: $S = I_{LO} + 2\mathfrak{Re}(E_{LO} \cdot E_{sig}) + I_{sig}$. I_{LO} is the homodyne of the LO, I_{sig} is the homodyne of the weak emitted signal, and $E_{LO} \cdot E_{sig}$ is the interference term that we want to extract. This heterodyne term contains all the response functions and phase dependency.

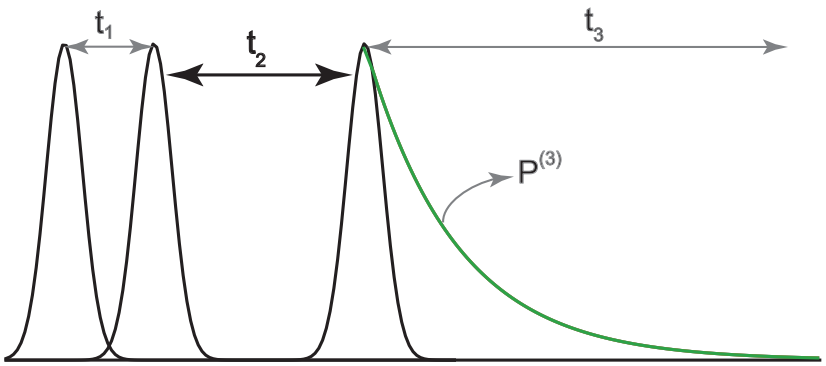

Figure 2.2: In a pump probe geometry, the first two pulses create a t_1 coherence. After a waiting time (t_2), the third pulse generates a 2D-IR signal ($P^{(3)}$) during the t_3 coherence.

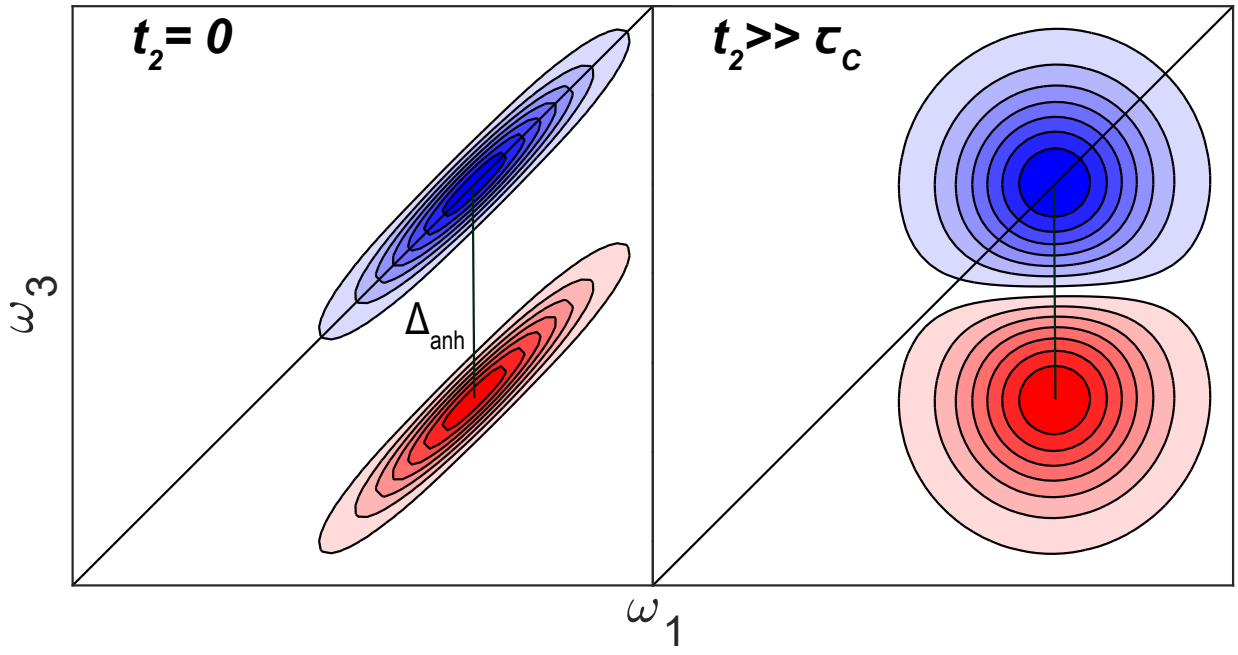

Figure 2.3: An example of spectral rounding in simulated spectra, where the increasing waiting time from zero to $t_2 \gg \tau_c$, where τ_c is the molecular correlation time, reveals the spectral diffusion. The spectra show a blue peak (0→1 transition) and the red peak is the anharmonically shifted 1→2 transition, where Δ_{anh} represents the anharmonicity.

In the frequency domain, the signal field (Figure 2.2) measured for a chromophore generates a diagonal peak with two components (Figure 2.3): the negative (blue) peak which corresponds to the 0–1 transition and the anharmonically shifted positive (red) peak corresponding to the 1-2 transition. After the initial excitation, the vibrational probe's frequencies are stretched along the diagonal indicating the correlation of the initial and final frequencies.

There are two limiting linewidths associated with a 2D-IR spectrum, the *homogeneous* and *inhomogeneous* widths. In the inhomogeneous limit (Figure 2.3 diagonal width of a peak), the molecules are observed as static during the initial perturbation by the 2D-IR pulses. With an increase in t_2 waiting time, the ensemble of molecules in the initial frequency start experiencing a change in that static frequency distribution due to alterations in the environment. On the other hand, every chromophore has an inherent bandwidth which is controlled by its vibrational lifetime and fast frequency fluctuations. This gives rise to the homogeneous limit (Figure 2.3 anti–diagonal width of a peak). In general, the lineshape of the 2D-IR spectrum depends on the system of interest and in some cases can be a combination of the inhomogeneous and homogeneous broadening.

Initially, the inhomogeneous distribution of frequencies is static (Figure 2.3: $t_2 = 0$). With increasing waiting time the molecules sample more of their surrounding. This changing initial environment around the chromophore produces fluctuations that manifest as frequency shift away from the diagonal (Figure 2.3: $t_2 >> \tau_c$). As a result, the spectrum with an increasing t_2 becomes round. This *spectral diffusion* process directly reports on the timescales at which the chromophores' surroundings evolve.

2.2 Molecular Vibrational Coupling

Vibrations can *couple*, and 2D-IR spectroscopy reports these couplings through cross-peaks in the spectra. A measure of this coupling gives us the distances and orientations of the individual molecules, furnishing structural information. Molecular models, spectral simulations, along with cross-peaks, in the spectrum provide comprehensive structural

knowledge.[29] In this section, we will first address how the anharmonic coupling between the vibrational states gives rise to a 2D-IR spectrum. After that, a local mode treatment of the Hamiltonian will demonstrate how 2D-IR spectroscopy better reflects the structural information than linear IR spectroscopy. Finally, we will also see how normal modes can give rise to a 2D-IR spectrum identical to the local mode depiction.

To track the geometry and reaction dynamics of a system, the concept of potential energy surfaces (PES) is useful (Figure 2.4 and Figure 2.4a). They map the energy of a molecule with respect to the atoms' $3N-6$ internal coordinates. Each of these internal coordinates can have different physical meanings, such as bond length, bond angle, torsion angle, phase angle, or plane angle. Utilizing these internal coordinates, it is possible to extract the PES of the molecule in $3N$ dimensions (a potential energy hyper-surface). We will focus on the example of two coupled modes. The 2D potential is

$$V(r_1, r_2) = V_a(r_1) + V_b(r_2) + \beta_{ab} r_1 r_2, \tag{11}$$

where β_{ab} is the coupling between the two vibrational potentials, V_a and V_b.

A Morse oscillator description of each individual potential is

$$V(r) = D(1 - e^{-ar})^2, \tag{12}$$

where r is the displacement about equilibrium in the respective bond coordinate, D_e is the equilibrium bond dissociation energy, and the a is the steepness of the potential and gives the harmonic force constant. In equation 11, the individual potentials $V_a(r_1)$ and $V_b(r_2)$, represent the 1D potential of each vibrational mode in the pair. A diagonalization of the Hamiltonian $H(r_1, r_2)$ leads to the eigenstates of this 2D potential. The equation 11 also introduces the idea of coupling between the local modes via the individual $V(r)$. Coupled local modes undergo frequency shifts, and, in PES terms the individual $V_a(r_1)$ and $V_b(r_2)$ change to a new $V(r_1, r_2)$. We will expand on these concepts shortly. The following subsections will show that we can deduce the same vibrational Hamiltonian for the system using both local and normal mode concepts.

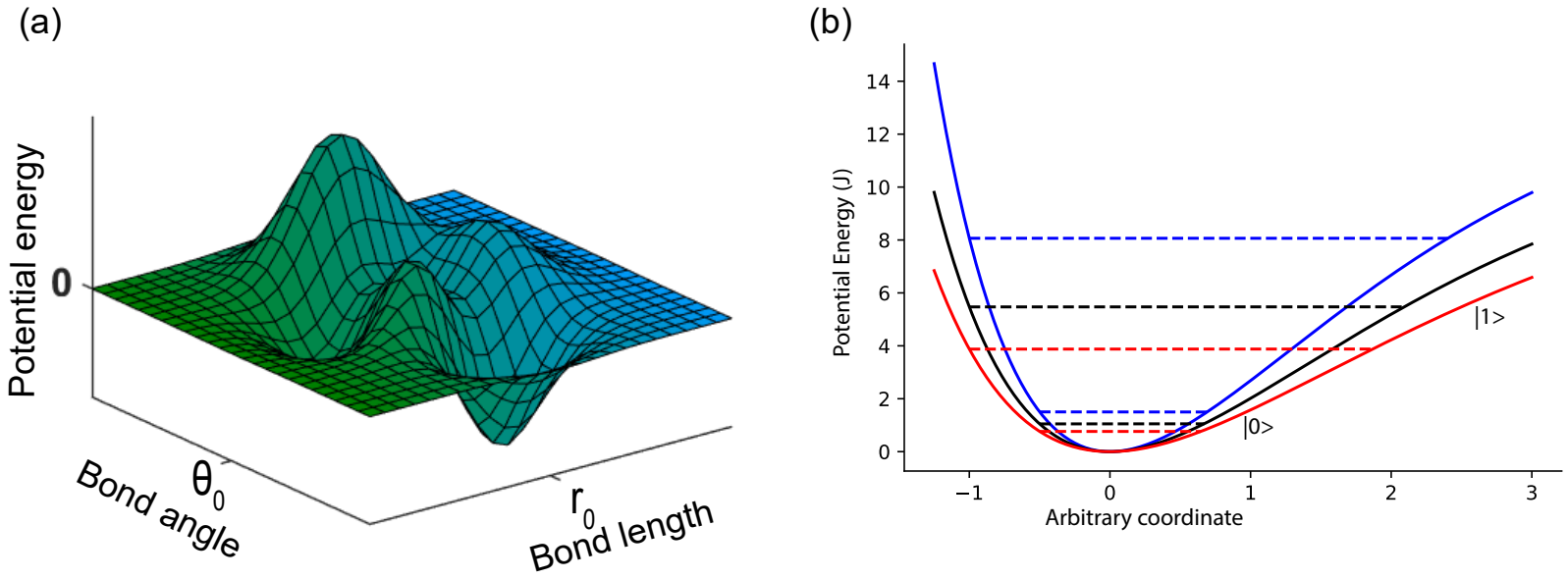

Figure 2.4: Interaction of the system with its surrounding encodes information about the structure and dynamics of the system. (a) A 2D-PES with bond length and bond angle coordinates. The equilibrium bond length and bond angle are set to r_0 and θ_0, respectively. (b) A 1D slice in a PES shows the stretch (red dashed line) and compression (blue dashed line) due to solvent interaction. In a bond length coordinate, the stretching and compression correspond to the red and blue shift of vibrational frequencies respectively.

2.2.1 A Local Mode Perspective

With the general introduction to the PES, we saw that an anharmonic description leads to a better approximation of any molecular system. Using the Morse oscillator model for each local mode vibration, we write the coupled Hamiltonian as

$$\hat{H} = \hbar\omega_i b_i^\dagger b_i + \hbar\omega_j b_j^\dagger b_j + \beta_{ij}(b_i^\dagger b_j + b_j^\dagger b_i) - \frac{\Delta}{2} b_i^\dagger b_i^\dagger b_i b_i - \frac{\Delta}{2} b_j^\dagger b_j^\dagger b_j b_j. \tag{13}$$

Here i and j are the two local modes, ω represents the frequency in wavenumber units, 'Δ' represents the local mode anharmonic shift and is proportional to the quartic expansion coefficients, when the PES is expanded by a Taylor expansion to get the higher order terms. Also, in equation 13 we are ignoring the zero-point energy and the further analyses in this section will continue doing that. The terms $b_{\mathrm{n}}^\dagger$ and b_{n} are the creation and annihilation operators, respectively. When we expand the Hamiltonian in the basis of $\{|ij\rangle\} = \{|00\rangle, |10\rangle, |01\rangle, |20\rangle, |02\rangle, |11\rangle\}$, we obtain a Hamiltonian where the first two block diagonals, the ground state and the first excited state, are identical to the Hamiltonian from a harmonic potential. The second excited state manifold, however, has local mode anharmonicity mixed inside it. The resultant Hamiltonian then undergoes a separate block diagonalization for each of its three manifolds. Under the assumption that $\omega_{\mathrm{a}} = \omega_{\mathrm{b}} = \omega$, the final Hamiltonian is

$$\hat{H} = \left(\begin{array}{c|cc|ccc} 0 & & & & & \\ \hline & \hbar\omega & \beta_{ij} & & & \\ & \beta_{ij} & \hbar\omega & & & \\ \hline & & & 2\hbar\omega - \Delta & 0 & \sqrt{2}\beta_{ij} \\ & & & 0 & 2\hbar\omega - \Delta & \sqrt{2}\beta_{ij} \\ & & & \sqrt{2}\beta_{ij} & \sqrt{2}\beta_{ij} & \hbar\omega + \hbar\omega \end{array}\right). \tag{14}$$

With only the harmonic potential we would not observe any 2D-IR spectrum, as the positive and negative peaks of a 2D-IR spectrum (Figure 2.3) would cancel each other. This local mode anharmonicity ($\Delta = \Delta_{local}$) mixing into the two–exciton state creates the diagonal and off–diagonal anharmonic shifts, which give rise to a 2D-IR spectrum.

2.2.2 Coupling between Two Local Modes

In this segment, we will connect the linear absorption spectrum to the geometry of two oscillators coupled to each other.[29,124] Time-independent perturbation in a system is expressed with an equation similar to equation 2,

$$\hat{H} = \hat{H}_0 + \hat{\beta}. \tag{15}$$

First we assume that the eigenfunctions representing the two oscillators are $|\phi_a\rangle$ and $|\phi_b\rangle$, with E_a $(= \hbar\omega_a)$ and E_b $(= \hbar\omega_b)$ eigenvalues, respectively. Our next assumption is that $\beta_{a,a}$ $= \beta_{b,b} = 0$. Now the $\hat{H}$ and $\beta_{a,b}$ can be expanded as

$$\hat{H} = \begin{pmatrix} E_a & \beta_{a,b} \\ \beta_{b,a} & E_b \end{pmatrix}, \tag{16a}$$

$$\beta_{a,b} = \beta^*_{b,a} = \langle\phi_a|\hat{\beta}|\phi_b\rangle. \tag{16b}$$

As the one exciton Hamiltonian in equation 16a couples the states, diagonalizing $\hat{H}$ will give us the mix between ϕ_a and ϕ_b; i.e, find the probabilities for the system to be in either state a or b. The mixed eigenfunction, $|\Phi\rangle$, can be expressed in the general form as a function of the two modes

$$|\Phi\rangle = c_a|\phi_a\rangle + c_b|\phi_b\rangle, \tag{17}$$

where c_a and c_b are the coefficients. Now the Schrödinger equation for the coupled states are

$$\begin{pmatrix} \hbar\omega_a & \beta_{a,b} \\ \beta_{b,a} & \hbar\omega_b \end{pmatrix} \begin{pmatrix} c_a \\ c_b \end{pmatrix} = E \begin{pmatrix} c_a \\ c_b \end{pmatrix}. \tag{18}$$

In order to get a non-zero solution to equation 18, we must solve

$$(\hbar\omega_a - E)(\hbar\omega_b - E) - \beta^2 = 0, \tag{19}$$

which gives the following eigenvalues, E_+ and E_-, for the mixed basis when solved:

$$E_\pm = \frac{\hbar\omega_a + \hbar\omega_b \pm \sqrt{(\hbar\omega_a - \hbar\omega_b)^2 + 4\beta^2}}{2}. \tag{20}$$

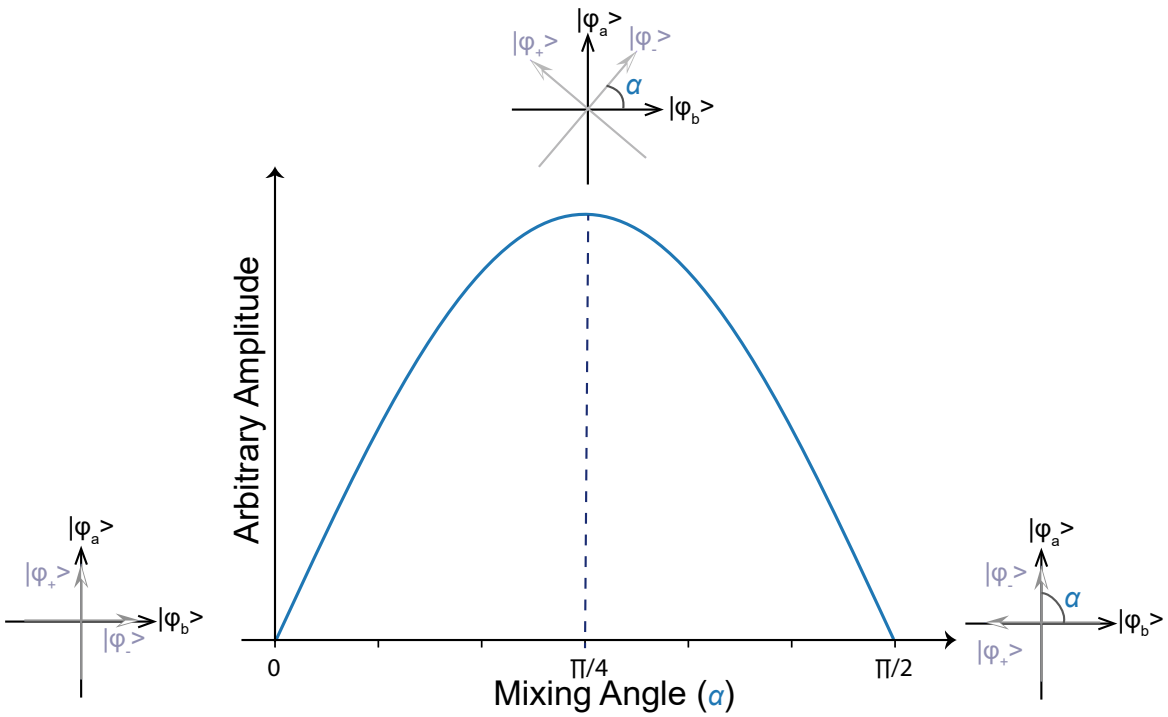

Figure 2.5: The mixing angle between the original and new bases is at a maximum at 45°. $\phi_{\pm}$ are the final bases, $\phi_{a/b}$ are the initial bases, and α is the mixing angle between the final and the initial bases.

The degree of mixing between the original modes is reflected in the eigenfunctions for the mixed states ($\Phi_\pm$) and the mixing angle (α).

$$\begin{pmatrix} |\Phi_+\rangle \\ |\Phi_-\rangle \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} |\phi_a\rangle \\ |\phi_b\rangle \end{pmatrix}, \tag{21}$$

where the mixing angle can be thought of as rotation from the space of uncoupled oscillators to the mixed basis via a basis *similarity transformation*. According to 16b, our Hamiltonian is symmetric. Therefore, the transformed Hamiltonian in the mixed basis is given by

$$H_\pm = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} H \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \tag{22}$$

where,

$$H_{\pm 11} = H_{11}\cos^2\alpha + H_{22}\sin^2\alpha + 2H_{12}\cos\alpha\sin\alpha \tag{23}$$

$$H_{\pm 22} = H_{22}\cos^2\alpha + H_{11}\sin^2\alpha - 2H_{12}\cos\alpha\sin\alpha \tag{24}$$

$$H_{\pm 12} = H_{12}\left(\cos^2\alpha - \sin^2\alpha\right) + (H_{22} - H_{11})\cos\alpha\sin\alpha \tag{25}$$

At a specific α value, the off-diagonal elements ($\beta_{\pm a,b}$) cease to exist in the new basis, and we can derive that mixing angle with the following relation

$$\begin{aligned} H_{12} &= H_{12}\left(\cos^2\alpha - sin^2\alpha\right) + (H_{22} - H_{11})\cos\alpha\sin\alpha \\ &= 0 \end{aligned} \tag{26}$$

$$H_{12}\left(\cos^2\alpha - sin^2\alpha\right) = (H_{22} - H_{11})\cos\alpha\sin\alpha \tag{27}$$

$$H_{12}\left(\cos 2\alpha\right) = (H_{22} - H_{11})\frac{1}{2}\sin 2\alpha \tag{28}$$

$$\tan 2\alpha = \frac{2H_{12}}{H_{11} - H_{22}}, \tag{29}$$

and in our equation 16a, this mixing angle (Figure 2.5) is,

$$0 < \alpha < \pi/2 \tag{30a}$$

$$\tan 2\alpha = \frac{2\beta}{\hbar\omega_b - \hbar\omega_a}. \tag{30b}$$

Similar to equations 21–30, the mixing between the transition dipole moments ($|\hat{\mu}_a\rangle$ and $|\hat{\mu}_b\rangle$) of the two modes ω_a and ω_b is

$$\begin{pmatrix} |\hat{\mu}_+\rangle \\ |\hat{\mu}_-\rangle \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} |\hat{\mu}_a\rangle \\ |\hat{\mu}_b\rangle \end{pmatrix}. \tag{31}$$

Equation 30 and 31 reflect that the two vibrational transitions' intensities ($|\hat{\mu}_+|^2$ and $|\hat{\mu}_-|^2$) are governed by the angle between the two dipoles, their coupling strength (β), and the energy of separation between the oscillators ($\hbar\omega_a$ and $\hbar\omega_b$). This is the direct connection between the molecule's geometry to the energy seen in the linear spectrum of two coupled absorbers.

From equation 30, we see that the mixing angle α can experience two extremes (Figure 2.5). In the weak coupling regime ($\alpha \sim 0°$ or $\alpha \sim 90°$), the modes are well separated but the coupling is small ($|V_{ab}| << |\hbar\omega_a - \hbar\omega_b|$). Hence, the vibrations are localized, indicating a small mixing angle (α), and the mixed state energies from equation 20 are

$$E_\pm \approx \hbar\omega_{a,b} \mp \frac{2\beta_{ab}^2}{\hbar\omega_a - \hbar\omega_b}. \tag{32}$$

In the strong coupling regime ($\alpha \sim 45°$), the coupling strength dominates, making $|V_{ab}| >> |\hbar\omega_a - \hbar\omega_b|$. The frequency separation between the modes become negligible compared to the coupling strength. Thus, the mixed state energies ($E_\pm$) in the strong coupling situation are perfectly delocalized with a mixing angle of $\alpha = \pi/4$. The split between the mixed state energies from equation 20 are given by

$$E_\pm \approx \frac{\hbar\omega_a + \hbar\omega_b}{2} \mp \beta_{ab}. \tag{33}$$

Transition Dipole Coupling (TDC) theory gives the simplest model of transition between two point dipoles. The transition dipole ($\hat{\mu}$), which is the transition between two unlike states $|1\rangle$ and $|2\rangle$ (equation 5), results from the radiation field perturbation (W in equation 3). In TDC theory, the TDM of a molecule undergoes a multipole expansion until the second order term.[125–127] In the weak coupling regime (equation 32), the effect of higher order terms on

the transition can be neglected, and the bilinear term survives. When $\hat{\mu}_1$ and $\hat{\mu}_2$ are the transition dipoles of the two vibrations and $\hat{r}_{12}$ is the displacement between the two, we have

$$\beta_{12} = \frac{1}{4\pi\epsilon_0}\left[\frac{\hat{\mu}_1 \cdot \hat{\mu}_2}{r_{12}^3} - 3\frac{(\hat{r}_{12} \cdot \hat{\mu}_1)(\hat{r}_{12} \cdot \hat{\mu}_2)}{r_{12}^5}\right]. \tag{34}$$

The two assumptions in TDC theory are : (1) there is electrostatic coupling between point dipoles and (2) through-bond phenomenon such as charge transfer and mechanical couplings, may be ignored. When the dipoles are in close proximity, TDC theory only makes a qualitative prediction of the magnitude and sign of couplings of the nearby modes disregarding couplings from covalently bound atoms.

2.2.3 Linear Spectrum of a Two Level System

The geometry of the dimer dictates the sign of coupling, β_{ij}. The orientation of the transition dipoles ($\hat{\mu}_1$ and $\hat{\mu}_2$) with respect to each other controls the sign of β_{ij}. This determines which of the two combinations appear in the high or low energy of the spectrum. There are five limiting geometries of the transition dipoles that give rise to the following coupling patterns in the linear spectrum. Look at Figure 2.6 as references to each case.

Case a: $\hat{\mu}_1$ is parallel to $\hat{\mu}_2$ in orientation. Equation 34 shows that β_{ij} is positive. In this case, the higher frequency solution is the symmetric eigenstate with transition dipole $\hat{\mu}_1 + \hat{\mu}_2$. As a result of the parallel orientation of the transition dipoles, this higher energy mixed state has most of the transition dipole strength. Comparatively, the lower frequency $\hat{\mu}_1 - \hat{\mu}_2$ antisymmetric eigenstate solution will have weak oscillator strength.

Case b: $\hat{\mu}_1$ is antiparallel to $\hat{\mu}_2$ in orientation. Here, the coupling will be negative, making the antisymmetric combination, $\hat{\mu}_1 - \hat{\mu}_2$, have the biggest oscillator strength and the higher frequency solution; however, the symmetric combination, $\hat{\mu}_1 + \hat{\mu}_2$, is the low energy solution with weak absorption.

Case c: $\hat{\mu}_1$ and $\hat{\mu}_2$ are collinear and point in the same direction. The low frequency transition is $\hat{\mu}_1 + \hat{\mu}_2$, has a high oscillator strength, whereas the high frequency transition, $\hat{\mu}_1 - \hat{\mu}_2$, is a weak transition.

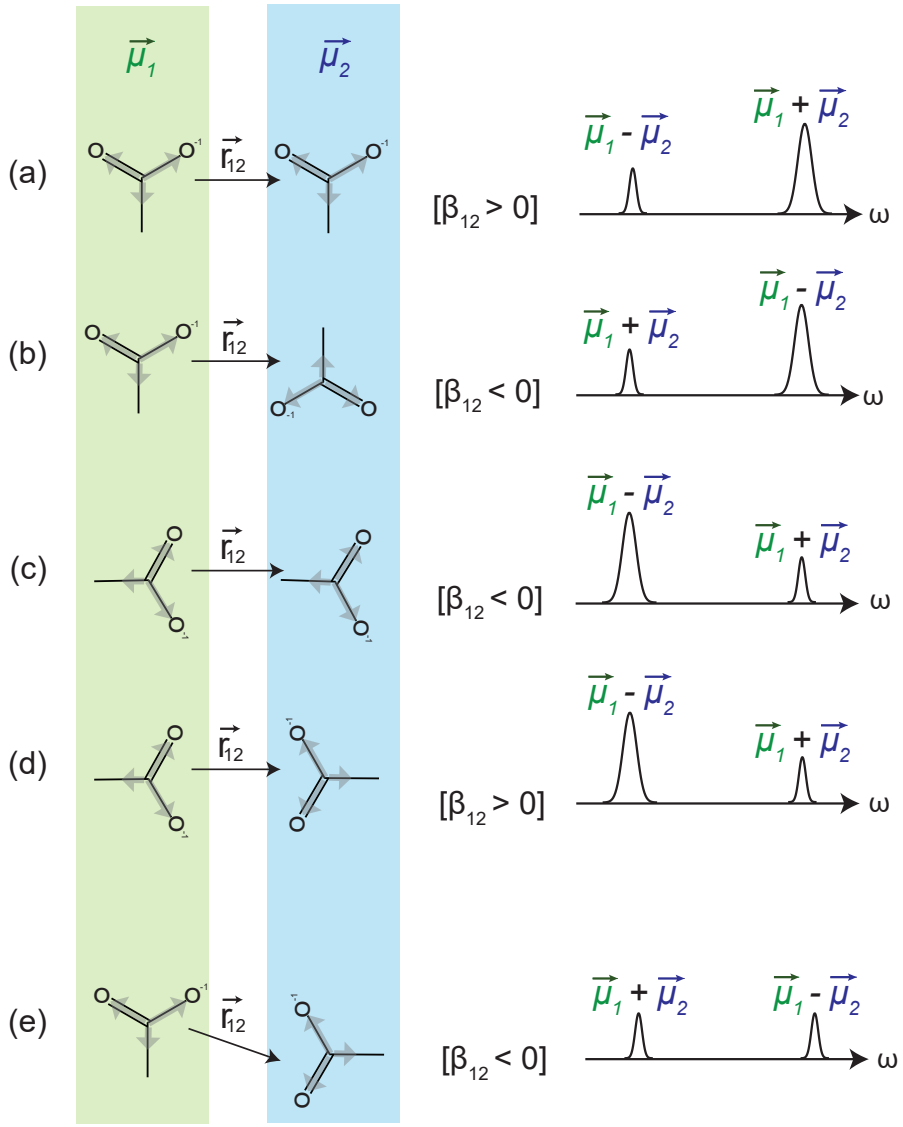

Figure 2.6: Different orientations of two coupled carboxylates produce similar linear spectra, making it difficult to extract geometry information. The light gray arrows on the carboxylate show its symmetric stretch mode. In the green block, every carboxylate orientation is assumed to have a transition dipole of $\vec{\mu}_1$, while in the blue block the transition dipoles for every orientation case are assumed to be $\vec{\mu}_2$. The separation between the two transition dipoles is $\vec{r}_{12}$ and the coupling constant from TDC theory is β_{12}.

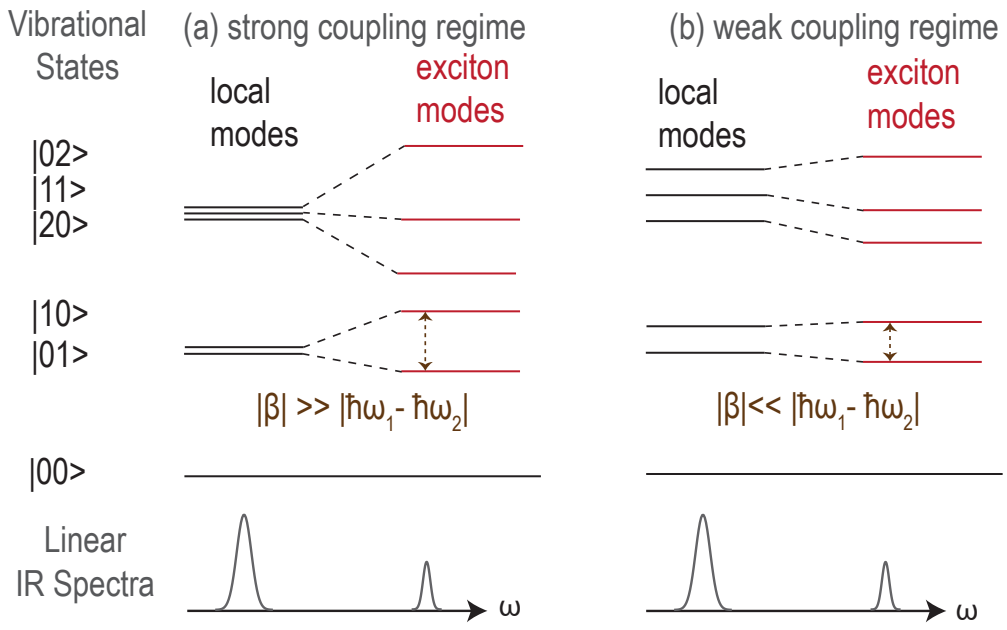

Figure 2.7: Strong coupling regime has a larger coupling constant than the weak coupling regime. The energy spacings differ between the two extremes, but we get the same linear infrared spectrum.

Case d: $\hat{\mu}_1$ and $\hat{\mu}_2$ are collinear and point in the opposite direction. In this case, the low energy transition is $\hat{\mu}_1 - \hat{\mu}_2$ with a high absorption, while the $\hat{\mu}_1 + \hat{\mu}_2$ is the high frequency solution with a low oscillator strength.

Case e: $\hat{\mu}_1$ and $\hat{\mu}_2$ are perpendicular to each other. When in this perpendicular orientation to each other, the resulting transitions are similar in intensities. That is the intensity of the $\hat{\mu}_1 + \hat{\mu}_2$ equals that of the $\hat{\mu}_1 - \hat{\mu}_2$ transition in the spectrum. The β_{ij}, is negative.

From cases a and b, we see that the two different orientations of the transition dipoles produce an identical linear spectrum. Similarly, cases c and d show us that the two other head-on orientations give the same linear vibrational spectrum. Therefore, cases (a–e) reveal that it is impossible to distinguish between the geometries of the transition dipoles from the linear spectrum alone. Also in the linear spectrum, the sum of intensities are the same in cases a–e, but there is a transfer of intensities between the transition dipoles due to coupling.

The energy gaps in the first two energy levels, above ground vibrational state for a two level system, in the strong and the weak coupling regimes differ. In the strong coupling regime (Figure 2.7a), the first and second vibrationally excited states are close in energy, ensuring a large coupling constant. This is reflected as the large energy gaps of the mixed or exciton modes. The weak coupling regime (Figure 2.7b), however, has a larger energy gap between the initial basis. This results in a low coupling constant reflected by the small energy gaps of the mixed bases. Linear infrared spectra are unable to capture these differences, even though the energy gaps between the vibrational levels show the obvious differences. In such cases, 2D-IR spectroscopy enables us to probe these three vibrational energy levels ($|0\rangle$, $|1\rangle$, and $|2\rangle$). Unlike linear IR spectroscopy, 2D-IR spectroscopy can clearly distinguish the strong from the weak coupling regime as we will see in the next subsection.

2.2.4 Molecular Structure from 2D-IR Spectrum of a Two Level System

Application of 2D-IR spectroscopy to the same coupled dimer studied above shows that the anharmonicity, Δ, mixes with the second exciton eigenstates (Figure 2.8). Mixing of Δ varies with the degree of delocalization. The cross-peaks thus generated distinguish between the two regimes. When strongly coupled, $|\beta_{ij}| \gg |\hbar\omega_2 - \hbar\omega_1|$, the vibration delocalizes over

the two local modes, causing greatly separated cross-peaks. In the weak coupling regime, $|\beta_{ij}| \ll |\hbar\omega_2 - \hbar\omega_1|$, the vibrations in these local modes are mostly localized. The cross-peak separations are now smaller than the respective diagonal bands. We will detail the role of anharmonicity in the normal mode discussion at the end of this section.

Cross-peaks in 2D-IR spectroscopy are easily observed, as the intensity is a function of $|\mu_i|^2|\mu_j|^2$. If a diagonal band ($|\mu^4|$ scaling) disappears, then the cross-peaks also vanish. The $|\mu^4|$ scaling of the diagonal peaks makes the bands more prominent in the 2D-IR spectrum than in a linear spectra, which only has $|\mu^2|$ intensity scaling. In both linear and 2D-IR spectra, the peak intensities depend on the orientation between the dipoles, the coupling strength, and the frequency gap between the oscillators. In 2D-IR spectroscopy, however, the sum of diagonal peak intensities changess with the change in coupling strength.

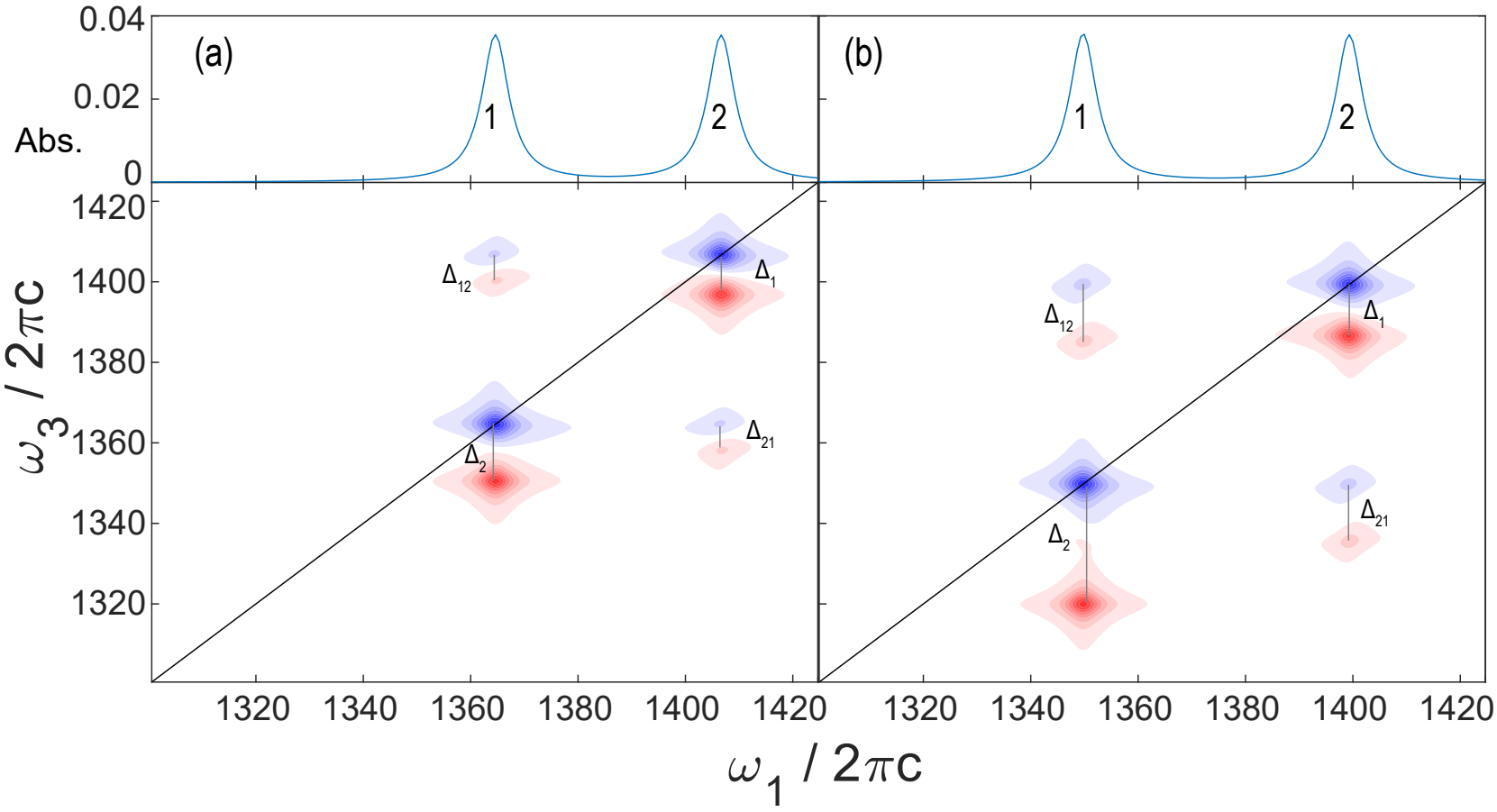

Figure 2.8: Anharmonicity in the cross-peaks is smaller when coupling is (a) weak versus (b) strong. Diagonal anharmonicities are Δ_1 and Δ_2 for peaks 1 and 2, respectively, while cross-peak anharmonicities are Δ_{12} and Δ_{21}. The top segment shows that the linear IR spectra for the two scenarios are similar.

There are two main factors facilitating the structure determination from a 2D-IR spectrum. The parameters extracted from the spectroscopy should allow us to calculate the PES. 2D-IR spectra give us the peak energies and intensities, which can be used to calculate couplings and the local mode terms (Figure 2.8). A 2D-IR spectrum provides the five terms $E_+, E_-, \Delta_{ii}, \Delta_{jj}$ and Δ_{ij}; however, we need $\omega_i, \omega_j, \Delta_i, \Delta_j$, and β_{ij} to extract the accurate frequencies. To decrease the number of unknowns required from a 2D-IR spectrum, we can utilize model compounds to measure the Δ_i and Δ_j terms. Polarization–controlled 2D-IR experiments can provide further information on the orientation, θ, of the transition dipoles. Overall, these molecular couplings in the 2D-IR spectra enable this spectroscopy to elucidate the system's structure.

Extracting coupling parameters between individual normal modes is a difficult venture, as infrared spectra have bandwidths associated with each peak. Coupling between the peaks can also be obscured by spectral congestion, making it impossible to extract the β_{ij} terms. Therefore, to generate a PES from a 2D-IR measurement, we should obtain the maximum number of parameters described above. The next step is to use these observables and predict a structure of the molecule, which requires a coupling model. TDC theory is the simplest model to consider these coupling interactions. There are more intricate coupling models that describe the PES more accurately than the TDC.[29,127] Some of these models involve calculating the PES using electronic structure theory or molecular dynamics simulations.

2.2.5 A Normal Mode Perspective

We demonstrated in section 2.2 how the local mode anharmonicity couples the second exciton manifold of the Hamiltonian and gives rise to the 2D-IR spectrum. Then throughout these previous subsections, we developed a local mode picture of the coupling between two vibrational modes and reinforced that the local mode coupling constants, β, provide a description of the Hamiltonian needed for the 2D-IR spectra. In this section, we will show that the same system in a normal mode representation will give us the same final Hamiltonian description required to produce a 2D-IR spectrum.

First we will solve the nuclear Schrödinger equation and get an in-depth understanding

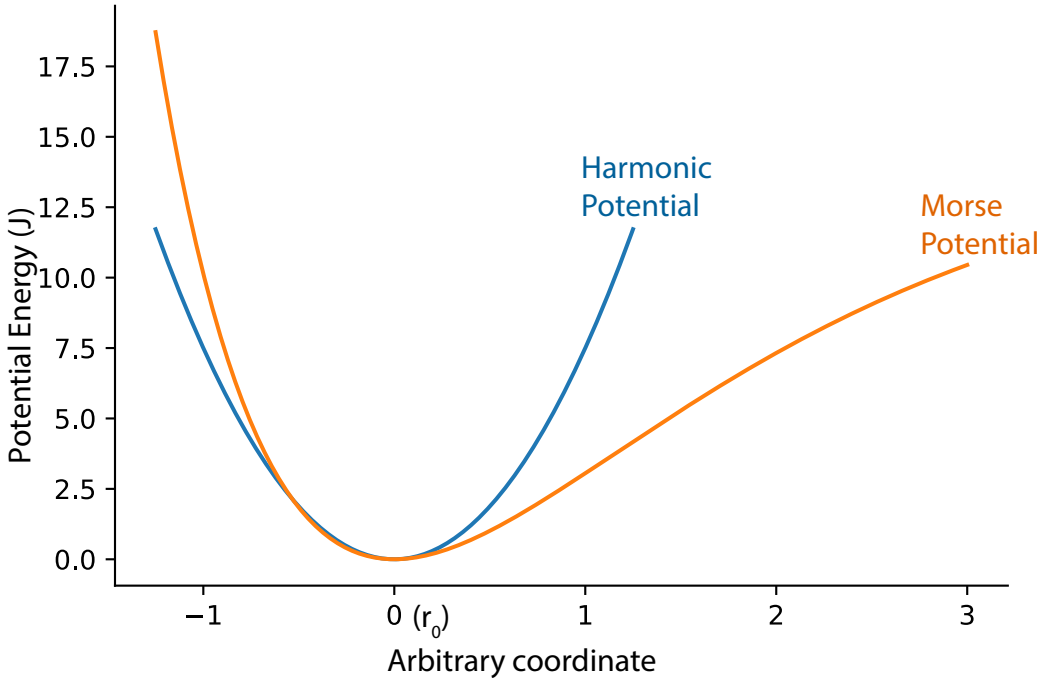

Figure 2.9: A comparison of both a harmonic and a Morse potential reveals that the latter includes anharmonicity, which can be extracted from the higher order terms in the potential energy expansion over a chosen coordinate. r_0 corresponds to the equilibrium '0' value and is the minimum bond length, when we are in bond length coordinates.

of the coupling process in the normal mode basis. Analysis of the problem here is similar to the treatment found in Mills and Robiette.[128] Solving a simple Morse oscillator for the vibrational energy terms first will give us the origin of the anharmonicity in the normal mode basis. Then we will extend those results and see how two normal modes can couple.

After adding a kinetic energy term to the Morse potential function (equation 12), we obtain the following Hamiltonian,

$$H = \frac{1}{2} g_{rr} \hat{p}_r^2 + D(1 - e^{-ar})^2. \tag{35}$$

In this Hamiltonian, the g_{rr}^2 term is the reduced mass and the $\hat{p}_r^2$ is the momentum operator. Expansion of the Morse potential about the equilibrium gives

$$V(r) = \frac{1}{2} f_{rr} r^2 + \frac{1}{6} f_{rrr} r^3 + \frac{1}{24} f_{rrrr} r^4 + ... \tag{36}$$

where the first term is quadratic, the second is cubic and the third is quartic. f_{rr}, f_{rrr}, and f_{rrr} are the quadratic, cubic, and quartic force constants, respectively. They are expressed as follows:

$$f_{rr} = \left.\frac{\partial^2 V}{\partial^2 r}\right|_0 = 2a^2 D_e. \tag{37a}$$

$$f_{rrr} = \left.\frac{\partial^3 V}{\partial^3 r}\right|_0 = -6a^3 D_e = -3a f_{rr}. \tag{37b}$$

$$f_{rrrr} = \left.\frac{\partial^4 V}{\partial^4 r}\right|_0 = 14a^4 D_e = 7a^2 f_{rr}. \tag{37c}$$

When the displacement about the equilibrium is small, the higher order terms are neglected and the cubic and quartic terms are treated as perturbations to the solutions of the zeroth order quadratic harmonic oscillator Hamiltonian.

$$H = H^{(0)} + H^{(1)}. \tag{38}$$

$$H^{(0)} = \frac{1}{2} g_{rr} \hat{p}_r^2 + \frac{1}{2} f_{rr} r^2 \tag{39a}$$

$$= \frac{1}{2} \hbar c \omega (\hat{p}^2 + q^2), \tag{39b}$$

$$H^{(1)} = \frac{1}{6} f_{rrr} r^3 + \frac{1}{24} f_{rrrr} r^4 \tag{39c}$$

$$= hc \left[\frac{1}{6} \phi_3 q^3 + \frac{1}{24} \phi_4 q^4 \right]. \tag{39d}$$

Here the Hamiltonian is expressed in terms of q and p, which are the dimensionless normal mode coordinate and its dimensionless momentum conjugate ($\hat{p} = -i \frac{\partial}{\partial q}$), respectively. This translates into the ladder operator formalism as $\hat{p}^2 + q^2 = 2BB^\dagger$. In the normal mode basis, ϕ_3 and ϕ_4 are the cubic and quartic anharmonic force constants in wavenumber units. ω is the oscillator frequency in wavenumbers. The vibrational eigenstates are obtained from the following equations,

$$q = \delta^{\frac{1}{2}} r, \delta = \frac{2\pi c \omega}{\hbar g_{rr}}. \tag{40a}$$

$$\omega = \frac{(g_{rr} f_{rr})^{\frac{1}{2}}}{2\pi c} = \frac{(2a^2 D_e g_{rr})^{\frac{1}{2}}}{2\pi c}. \tag{40b}$$

$$\phi_3 = \frac{f_{rrr}}{hc\delta^{\frac{3}{2}}}, \phi_4 = \frac{f_{rrrr}}{hc\delta^2}. \tag{40c}$$

Using second order perturbation theory to treat the cubic term and a first order perturbation theory for the quartic term, we get the eigenstates as

$$E = \hbar\omega_1 \left(n + \frac{1}{2} \right) + \Delta \left(n + \frac{1}{2} \right)^2. \tag{41}$$

The anharmonicity constant

$$\Delta = \frac{\phi_4}{16} - \frac{5\phi_3^2}{48\omega}, \tag{42}$$

is obtained from the above perturbation treatment of the cubic and quartic terms and establishes the importance of these two terms in the origin of anharmonicity. Substituting

equations 40 and 37 in equation 42 gives us the anharmonicity in terms of the Morse parameters,

$$\Delta = -\frac{a^2\hbar^2 g_{rr}}{2hc}. \tag{43}$$

In the next example, the PES is expanded in the basis of two normal coordinates, q_1 and q_2.

$$V(q_1, q_2) = \frac{1}{2}\sum_i \frac{\partial^2 V}{\partial^2 q_i}\bigg|_0 q_i^2 + \frac{1}{6}\sum_{ijk} \frac{\partial^3 V}{\partial q_i q_j q_k}\bigg|_0 q_i q_j q_k + \frac{1}{24}\sum_i \frac{\partial^4 V}{\partial q_i q_j q_k q_l}\bigg|_0 q_i q_j q_k q_l + ... \tag{44}$$

Similar to equation 36, the first term in equation 44 is quadratic, the second term is cubic, and the third term is quartic. From equation 42, we see that anharmonicity arises from the cubic and quartic terms, with a quadratic dependence to the former and a linear dependence to the latter. In ladder operator formalism (similar to Section 2.2.1), if we apply $B_i^\dagger B_j^\dagger B_j B_j^\dagger$ on the diagonal states such as $|00\rangle$ and $|11\rangle$, we extract the anharmonicity on one coordinate. Moreover, in these same $q_i q_j q_k q_l$ terms there are couplings between the $|10\rangle$ and $|10\rangle$ states, which can be represented by $B_i B_i^\dagger B_j B_j^\dagger$. $q_i q_j q_k$ includes terms that couple different states such as $|00\rangle$ with $|01\rangle$ and $|01\rangle$ with $|11\rangle$. We can use $B_i B_i^\dagger B_i^\dagger$ and $B_i^\dagger B_i B_j$ to extract those couplings. This creation and annihilation operator treatment gives us a comprehensive set of all the couplings in the Hamiltonian, including those inside and outside of the block diagonals. These can be schematically represented as,

$$\hat{H} = \left(\begin{array}{c|c|c} B_i B_i^\dagger B_j B_j^\dagger & h.c. & h.c. \\ \hline B_i^\dagger B_j B_j^\dagger & B_i B_i^\dagger B_j B_j^\dagger & \\ & B_i^\dagger B_j B_j B_j^\dagger & h.c. \\ \hline B_i B_i^\dagger B_j^\dagger B_j^\dagger & & B_i B_i^\dagger B_j B_j^\dagger \\ B_i^\dagger B_j^\dagger B_j B_j^\dagger & B_i^\dagger B_j B_j^\dagger & B_i^\dagger B_j B_j B_j^\dagger \\ & & B_i B_i B_j^\dagger B_j^\dagger \end{array}\right), \tag{45}$$

where, 'h.c.' represent the Hermitian conjugates.

To remove the coupling elements between the blocks in the basis of two normal modes, we use degenerate perturbation theory and block diagonalize the Hamiltonian in n, where the total excitation n $= n_1 + n_2$ and the total basis functions are $n + 1$. *Dunham expansion*, obtained from degenerate perturbation theory, is often used to calculate normal mode energies.

The following equations give the non-zero matrix elements, where ω_i and Δ_{ii} represent the normal mode frequencies and anharmonicites, respectively.

$$\begin{aligned} E_{n_1,n_2} = \langle n_1 n_2|H|n_1 n_2\rangle = \hbar\omega_1(n_1+\frac{1}{2}) + \Delta_{11}(n_1+\frac{1}{2})^2 + \\ \hbar\omega_2(n_2+\frac{1}{2}) + \Delta_{22}(n_2+\frac{1}{2})^2 + \Delta_{12}(n_1+\frac{1}{2})(n_2+\frac{1}{2}), \end{aligned} \tag{46}$$

$$\langle n_1+2, n_2-2|H|n_1 n_2\rangle = \frac{1}{2}\gamma[(n_1+1)(n_1+2)n_2(n_2-1)]^{\frac{1}{2}}, \tag{47a}$$

$$\langle n_1-2, n_2+2|H|n_1 n_2\rangle = \frac{1}{2}\gamma[n_1(n_1-1)(n_2+1)(n_2+2)]^{\frac{1}{2}}. \tag{47b}$$

The Δ_{ij} term corresponds to the intermode or off-diagonal anharmonicity, that can be directly read off from a 2D-IR spectrum from the splitting between the cross-peak pairs. The n_i is the quantum number of a normal mode i and γ is the *Darling-Dennison* term that couples the nearly degenerate states. Next we will go through an example of a 0-1 transition of the normal mode ω_1.

$$\begin{aligned} \nu^{(1)}_{0\to1} = E_{1,n_2} - E_{0,n_2} = \langle 1n_2|H|1n_2\rangle - \langle 0n_2|H|0n_2\rangle \\ = \hbar\omega_1 + 2\Delta_{11} + \Delta_{12}(n_2+\frac{1}{2}) = \nu^{(1)} + \Delta_{12}n_2, \end{aligned} \tag{48}$$

where the term $\nu^{(1)} = \hbar\omega_1 + 2\Delta_{11} + \frac{\Delta_{12}}{2}$. So from equation 48, we see that when mode $\nu^{(2)}$ is in the ground state ($= |0\rangle$), then $\nu^{(1)}_{0\to1} = \nu^{(1)}$ has a slight frequency shift due to the intermode coupling term ($2\Delta_{11}$) in the expression. When the normal mode $\nu^{(2)} \neq 0$, then the final expression for $\nu^{(1)}_{0\to1}$ is given by equation 48, where the off-diagonal anharmonic shift (Δ_{12}) participates significantly in the coupling of the two modes. Using Dunham expansion, a similar expression for the transition $\nu^{(1)}_{1\to2} = \hbar\omega_1 + 4\Delta_{11} + \Delta_{12}(n_2+\frac{1}{2})$ can be obtained.

A parallel between the normal modes and the local modes is drawn by imposing the following relations between the parameters in equations 46 and 47: $\omega_{local} = \omega_{01} - 2\Delta_{local}$.

$$\omega_1 = \omega_{local} + \beta, \omega_2 = \omega_{local} - \beta, \tag{49a}$$

$$\Delta_{11} = \Delta_{22} = \frac{\Delta_{12}}{4} = \frac{\Delta_{local}}{2} \tag{49b}$$

$$\gamma = \Delta_{local}. \tag{49c}$$

In the normal mode Hamiltonian, there are still off-diagonal terms in the second block manifold that couple nearly-degenerate states. Degenerate perturbation theory is not applicable to these $B_i B_i B_j^\dagger B_j^\dagger$ terms. *Darling-Dennison* resonance Hamiltonian[129] can model these off-diagonal terms. After applying the Darling-Dennison coupling terms and subtracting the zero-point energy, the final form of the Hamiltonian in this normal mode basis becomes

$$\hat{H} = \left(\begin{array}{c|cc|cccc} 0 & & & & & & \\ \hline & \hbar\omega - \beta & & & & & \\ & & \hbar\omega - \beta & & & & \\ \hline & & & 2\hbar\omega - 2\beta - \frac{\Delta}{2} & -\frac{\Delta}{2} & & \\ & & & -\frac{\Delta}{2} & 2\hbar\omega - 2\beta - \frac{\Delta}{2} & & \\ & & & & & 2\hbar\omega - \Delta & \end{array}\right). \tag{50}$$

Therefore, the Darling-Dennison coupling terms give rise to the off-diagonal elements in the second block manifold. A complete diagonalization of this Hamiltonian yields the identical eigenstates as in the local mode Hamiltonian treatment (equation 14). Both local mode and normal modes have described the same molecular eigenstates. The Morse oscillator local modes are coupled by β, whereas the normal modes are coupled by Δ coming from Dunham perturbation and Darling-Dennison terms. To choose one description from the other, we check if β or Δ is acting as a perturbation. When $\beta \gg \Delta$, Δ is the perturbation and the normal mode picture is a better description of the molecular eigenstates. If $\Delta \gg \beta$, where β acts as a local mode perturbation, then the local modes closely represent the molecular eigenstates. If the relations in the equations 49 hold true, then both local and normal modes will equally depict the molecular eigenstates; however, in reality only one basis is the closest representation of the true molecular energies of the system. During modeling of the system, the basis closest to the true molecular eigenstate (local or normal) is chosen to minimize the error in representation.

Sections 2.1 and 2.2 introduce the general concepts of 2D-IR spectroscopy and how to get structural information from it. Section 2.2 specifically addresses the importance of both local and normal mode couplings in this spectroscopy. These couplings can be obtained

by performing *ab initio* calculations. Using the vibrations, transition dipoles, and the IR intensities from electronic structure theory, one can calculate the 2D-IR spectra. This yields a better understanding of the experimental system and also helps improve calculation parameters for a better comparison with the experiment.

To calculate the 2D-IR spectra for a molecule, we can either use the harmonic vibrations or the anharmonic frequencies. For large molecules, such as EDTA used in chapter 4, an anharmonic vibrational frequency calculation is time-consuming so a normal mode calculation is preferred. In the case of small molecules, such as H_2O, one can calculate anharmonic frequencies in a time-effective manner. New and improved computational methods can calculate the local mode coupling constants from the vibrational Hamiltonian of the system. To comprehend how some of these molecular properties are calculated, brief descriptions of electronic structure theory and density functional theory is covered in the next section.

2.3 Electronic Structure Calculation

This section will first address why we want to calculate molecular properties. Then some widely used methods are compared. Basics of density functional theory follows next. All these discussions will lead to a better understanding of the electronic structure calculations performed in chapter 4. This section's discourse is based on the textbooks by Frank Jensen[130] and by Sholl and Steckel.[131]

2.3.1 Calculating Molecular Properties

Computational chemistry is a burgeoning discipline that is capable of probing the chemical characteristics of molecules. With rapid growth in computational power, advent of supercomputers, and advances in computer algorithms and programs, chemists rely more than ever on computational quantum chemistry to grasp the intricate behavior of molecules. Before taking advantage of any computational calculation, we must correctly describe our chemical system.

For a complete system description we must: (1) First define the fundamental units of the system and their number, which is constrained by the desired chemical property. Particles can be atoms, molecules, or entire protein structures. In the process, we lose finer information about smaller details within our particle of choice; (2) Have the particles' initial positions and their velocities. 'Starting conditions' define the first instances of our chosen system. It can be a portion of the whole sample space. (3) Examine the interactions that govern the particles' behavior and the forces acting between them. (4) Understand how the system evolves with time. In (3) and (4), 'interaction' between particles and 'dynamical equations' together solve the evolution and behavior of the system with time. Choosing the forces that describe each interaction depends on the system under scrutiny.

After establishing these fundamental chemical aspects, we apply them as input parameters in computational chemistry calculations to solve for molecular properties, such as equilibrium and transition state chemical structures, dipole and quadrupole moments and polarizabilities, IR and Raman vibrational frequencies, NMR spectra, UV-Vis spectra, reaction rates, and themochemical attributes.

Understanding the molecular properties of a system requires insight into the changes in the energy levels of the system. The atom in a molecule has electrons and the nucleus, of which the former is 10^3 times lighter than the latter. Thus, electron motion is faster than the nucleus and this is the basis of the Born-Oppenheimer approximation. With the Born-Oppenheimer approximation, we can separate out the electron motion from the nucleus and develop the adiabatic PES of atoms. The simplest non-relativistic time-independent Schrödinger equation that we are familiar with (equation 51) has an exact analytical solution only for a two-body system with one nucleus and one electron. With many-body complex molecular systems, getting an exact solution is difficult. Due to the Born-Oppenheimer approximation, we can first solve the electronic problem for a set of nuclei. After solving this electronic energy hyper-surface over a set of nuclear coordinates, it is easier to solve for the nuclear motion on the potential energy hyper-surface.

$$H\Psi = E\Psi, \tag{51}$$

and when the H expression is fully expanded we have

$$H = \Bigg[-\frac{\hbar^2}{2}\sum_{\alpha=1}^{N}\frac{1}{m_\alpha}\nabla_\alpha^2 - \frac{\hbar^2}{2m_e}\sum_{i=1}^{N}\nabla_i^2 + \sum_{\alpha=1}^{N}\sum_{\beta<\alpha}\frac{Z_\alpha Z_\beta e'^2}{r_{\alpha\beta}} - \sum_{\alpha=1}^{N}\sum_{i}\frac{Z_\alpha e'^2}{r_{i\alpha}} + \sum_{i=1}^{N}\sum_{j<1}\frac{e'^2}{r_{ij}}\Bigg]. \quad (52)$$

All the α and β terms pertain to the nuclei, while all the i and j terms represent the electron, m describes mass, Z is the atomic number, e' is the electronic charge, and r is the distance between two particles. From left to right in equation 52, the first and second terms are nuclear and electronic kinetic energies respectively, the third term is nuclear–nuclear repulsion, the fourth term is nuclear–electronic attraction, and the last term is electron–electron repulsion.

After the Born-Oppenheimer approximation, the relevant electronic Hamiltonian is

$$H = \Bigg[\frac{\hbar^2}{2m}\sum_{i=1}^{N}\nabla_i^2 + \sum_{i=1}^{N}V(r_i) + \sum_{i=1}^{N}\sum_{j<1}U(r_i r_j)\Bigg], \quad (53)$$

where the first term is the kinetic energy of each electron, the second term represents the interaction between each electron and the collection of atomic nuclei, and the third term is the interaction between all the electrons. The last two terms together make the electronic potential energy, where $V(r_i)$ and $U(r_i r_j)$ in equation 53 are equal to the nuclear–electronic attraction and electron–electron repulsion terms in equation 52, respectively. In equation 51, E is the electronic energy of the ground state and $\Psi = \psi(r_1, r_2, ..., r_N)$ reflects that the wavefunction is a function of all the electron coordinates. We can convert Ψ into the product of the individual ψ_i electron wavefunctions,

$$\Psi = \prod_{i=1}^{N}\psi_i(r), \quad (54)$$

where N is much greater than the number of nuclei. This increases the dimensionality of the problem considerably making the solution process time–expensive. For example, in this formulation of full wave functions, the linear molecule CO_2 would be a 66 dimension problem.

In Hamiltonian (equation 53), the electron-electron interaction is the most complicated and difficult to solve term, as the $\sum_{i=1}^{N}\sum_{j<1}U(r_i r_j)$ term contains all interactions of every electron with all the other electrons in the system. Therefore, this analysis shows that the Schrödinger equation is a 'many-body problem' including all these individual interactions.

2.3.2 Comparing Computational Methods

As mentioned earlier, the choice of the system determines the details that can be expected from a PES calculation result. To solve a PES, we can use force field methods which use atoms as building blocks, dynamical equations from classical mechanics, and ignore the electronic PES. Here, the explicitly provided fitting parameters directly correlate with the physical experiment. If, however, we want a full analysis of the electronic distribution of the system, we can resort to methods such as Hartree-Fock (HF), Møller-Plesset perturbation (MP2, MP4) theory, and coupled cluster (CC). The methods that do not use any experimental reference as a starting point are the *ab initio* methods, while the ones that take experimental parameters into account are the *semi-empirical* methods.

The *ab initio* methods include HF theory, which solves the many-body Schrödinger equation by approximating the electron-electron correlation term as an average. Density functional theory (DFT), which we will discuss shortly, also falls under such *ab initio* methods. The *semi-empirical* methods are modeled from HF theory, where no more than two nuclei are used to construct the significant equations in the theory. The rest of the integrals are then used as fitting parameters for experimental data, such as, molecular energies and geometries. *Semi-empirical* methods are limited to systems for which experimental parameters are known.

DFT is an independent particle model and compares with HF theory. A disadvantage of DFT compared to other *ab initio* methods, however, is the absence of a systematic approach to improving an existing result to get an exact solution. How DFT circumvents this problem will be discussed later in this section.

Two things of importance when choosing a computational method and solving a chemical question are the computational cost and the accuracy of the method for that particular problem. The immediate parameter that contributes to the computational scaling is the size of the chemical system. When the basis set size is fixed, the relative scaling (denoted in parenthesis, where the number of particles is N) between some of the methods are

$$\text{DFT}(N^3) \leq \text{HF}(N^{3-4}) < \text{MP2} = \text{CC2}(N^5) < \text{CCSD} = \text{MP4}(N^6) < \text{Full CI}(N!), \quad (55)$$

where 'CCSD' is Coupled Cluster with Singles and Doubles and 'Full CI' is Full Configurational Interaction.

To determine the system's energy accurately for a given medium sized basis set, the following order is observed in these methods:

$$\text{HF} << \text{MP2} < \text{CISD} < \text{MP4(SDQ)} \sim \text{CCSD} < \text{MP4} < \text{CCSD(T)}. \tag{56}$$

The order in equation 56 is only valid when the one determinant methods are compared. The antisymmetry of the total electronic wavefunction to the exchange of two electrons is mathematically constructed inside *Slater Determinants (SDs)*. In the SD, the rows represent the electron coordinates while the columns are the orbitals. Utilizing such a single SD as a trial wave function, one can derive the set of HF equations and minimize for energy. Multiple Slater determinants improve the energy obtained from HF theory. In multi determinant methods, the full Slater determinant space represents the total wavefunction. Unlike the one electron basis where the basis set determines its size, a many electron basis, size is determined by the number of SDs. In the equation 56 scheme, HF is a 'good' zeroth-order description of the system. With increasing multi determinant characters, the CC methods, with their 'infinite' order descriptors, perform better than the perturbation methods. DFT falls somewhere in between these methods in accuracy, depending on the chemical problem being addressed. For small molecules, DFT and HF theory have comparable outputs. For DFT in larger systems, however, improving the basis sets and the *functional* can lead to an accurate description of the system.

2.3.3 Density Functional Theory

Equation 54 reflects the probability of N-electrons being at a particular set of coordinates i.e, $\psi^*(r_1, r_2, ..., r_N)\psi(r_1, r_2, ..., r_N)$, where ψ^* is the complex conjugate of ψ. In this wavefunction description, the system size is defined by the total number of electrons (N) and it has $4N$ variables, including the three spatial coordinates and one spin coordinate for each electron. This probability density expression is related to the density of electrons at a particular position, $n(r)$, in space by

$$n(r) = 2\sum_i \psi_i^*(r)\psi_i(r). \tag{57}$$

The sum in equation 57 is over all the individual wavefunctions that are occupied by electrons. So $\psi_i^*\psi_i$ is the probability of finding an electron at 'r'. The prefix '2' comes from the electron spin, where the Pauli exclusion principle states that two electrons with the same spatial address must have different spins, and we will expand on this later in the section. After applying the electron density formulation to the original wavefunction problem, the dimensionality is reduced from '$3N$' down to just '3'. This drastic reduction in dimension in essence captures the popularity and utility of density functional theory. Now a daunting dimension problem can be better handled in three dimensions.

Density functional theory stands on two fundamental theorems:

1. *The ground state energy from the Schrödinger equation is a unique functional of the electron density.*
2. *The electron density that minimizes the energy of the overall functional is the true electron density corresponding to the full solution of the Schrödinger equation.*

These two theorems were both proven by Kohn and Hohenberg. Later, Kohn and Sham derived the necessary equations associated with the theorems. The first theorem alludes to the presence of a complete electron density description for the ground state electron wavefunction. We can use that electron density to determine the properties of the system in the ground state. Therefore, with the first theorem, we can solve the system's ground state energy in just three coordinate dimensions. Due to this first theorem, we know there is a way to extract the ground state energy of the system; however, we do not know how to determine that *functional*.

First, let us understand what precisely the functional is doing in this theory. A 'functional' takes a function as an input parameter to calculate an output number. As for example, let's say $F[f] = \int f(x)dx$, so $F[f]$ is a value and $f(x)$ is any function we want to evaluate. Say we use $f(x) = x^2 + 1$, so we will get $F[f] = \frac{2}{3}$. In the first Hohenberg-Kohn theorem, the ground state energy, $E[n(r)]$, is the number that gets evaluated from the electron density function, $n(r)$. Therefore $E[n(r)]$, is the functional for the ground state energy that we approximate by utilizing the second theorem.

Only if the *absolute* form of the functional was known, the second theorem states that the

electron density can be completely calculated by minimizing the energy from the functional. The Schrödinger equation is a many-body problem and we will now expand upon that to develop DFT. We write equation 53 as

$$E[\psi_i] = E_{known}[\psi_i] + E_{XC}[\psi_i]. \tag{58}$$

The first term in equation 58 in electron density terms is

$$E_{known}[\psi_i] = \frac{h^2}{m}\sum_{i=1}\int \psi_i^*\nabla_i^2\psi_i d^3r + \int V(r)n(r)d^3r + \frac{e^2}{2}\int\int \frac{n(r)n(r')}{|r-r'|}d^3rd^3r' + E_{ion}. \tag{59}$$

In equation 59, the first term is the electron kinetic energy, the second term is the Coulomb interaction between the electrons and nuclei, the third term is the Coulomb interaction between the pairs of electrons, and the last term is the Coulomb interaction between pairs of nuclei.

Equation 58 has another term, $E_{XC}[\psi_i]$, the *exchange–correlational functional*. This term includes all the quantum mechanical effects that are not addressed by the E_{known} terms. We still need to know how to evaluate the minimum of $E[\psi_i]$, which is obtained from the total energy functional. Kohn and Sham laid out this groundwork. When the electronic equations are expressed to solve a set of single–electron wavefunction equations for each electron, then we converge towards the complete solution of the Schrödinger equation. The Kohn–Sham equations take the form

$$\left[\frac{h^2}{2m}\nabla_i^2 + V(r) + V_H(r) + V_{XC}(r)\right]\psi_i(r) = E_i\psi_i(r). \tag{60}$$

The above equation looks similar to equation 59 but it lacks the summation terms, has single electron wavefunctions, and is defined in the three Cartesian coordinate system. Potential $V(r)$ in equation 60, the same as in equation 59, describes the interaction between an electron and a set of atomic nuclei. $V_H(r)$ is called the Hartree potential and goes by $V_H(r) = e^2\int\frac{n(r')}{|r-r'|}d^3r'$. It defines the Coulomb repulsion between the electron of interest in one Kohn–Sham subset equation and the total electron density of the whole system. There is a self interaction term that originates in the $V_H(r)$ term and it is due to the electron being a part of that total electron density. This unphysical self–interaction term must be corrected for, and DFT uses the $V_{XC}(r)$ term to do that. We will discuss shortly.

Finally, the last potential, $V_{XC}(r)$, accounts for the exchange and correlation interactions for a single electron equation, and it is called the *functional derivative* of the *exchange–correlation energy.*

$$V_{XC}(r) = \frac{\partial E_{XC}}{\partial n(r)}. \tag{61}$$

The next step is to solve these Kohn–Sham (KS) equations. First, the Hartree potential has to be known, in which case we should know the electron density. The electron density in turn comes from the single electron wavefunctions, which are again a product of solving the KS equations. In order to get past this circular argument and solve the KS equations, the following iterative approach is used:

1. Define a starting trial electron density, $n(r)$.
2. Use the trial electron density to get the single particle wavefunctions ($\psi_i(r)$), which in turn helps in solving the KS equations.
3. Evaluate the electron density with $\psi_i(r)$ from the previous step, $n_{KS} = 2\sum_i \psi_i^*(r)\psi_i(r)$.
4. Compare the trial $n(r)$ with the calculated $n_{KS}(r)$. When the two electron densities match, we have reached the ground state electron density that can give us the total energy. If they do not match, then $n(r)$ is first updated to the calculated ($n_{KS}(r)$) and steps 2–4 are repeated.

Following this iterative approach leads to a solution of the KS equations that is *self–consistent.*

2.3.4 Exchange Correlation Functional

Knowing the exchange–correlation functional exactly is daunting, as its modeling requires an infinite set of functions. The exact form of the functional is only known for the uniform electron gas. For every other system, the $V_{XC}(r)$ functional is approximated. The electron gas problem forms the basis of the local density approximation (LDA), which is the starting point of the systematic approach of improving these KS functionals.

One might ask how exchange and correlation energies contribute to the electron density problem. The exchange energy originates from the interaction between two electrons. Coulomb's equation is the classical representation of the energy from electron repulsion due to their similar charges. Applying quantum mechanics leads to modification of the

Coulomb equation to account for the $\frac{1}{2}$ spins of electrons. By the Pauli Exclusion principle, two fermions cannot both have the same spatial address and spin. Therefore, the total wavefunction must be antisymmetric to interchange between any two such particles, forming the basis of exchange energy. Exchange energy is the quantum mechanical correction to the Coulomb repulsion and accounts for the electron self–repulsion.

Even when the exchange energy is defined by the wave mechanics equations, the interpretations of these equations in KS theory utilizes the electron density. So, this gives rise to non–local potentials, which make an exchange potential at a given point erroneously dependent on the electron density at a distant point. Correlation energy, in wave mechanics, is defined as the difference between the exact energy and the corresponding HF value.

Considering the distances between two electrons, both exchange and correlation energies can be *long–range* or *short–range*. The long–range correlation, also called *static correlation*, cancels the short–range exchange energy in wave mechanics and needs to be built into the E_{XC} functional used in DFT, because DFT relies on electron density at a point and its immediate neighborhood. The short–range correlation is also called *dynamical correlation*. Ultimately, the effect of the static versus dynamic correlation is microscopically indistinguishable when the approximations become increasingly exact. For example, compared to DFT, the Full CI calculation does not distinguish between these two categories.

A physically intuitive picture of the exchange and correlation energies comes from using the concept of *exchange* or *correlation holes*. The exchange term, already derived in Hartree–Fock theory, is incorporated into DFT and the additional dynamical effect from electrons avoiding each other gives rise to the correlation energy. If we assume that electrons lack both charge and spin, then the probability of finding an electron (e_1) at any defined position is independent of the presence of a second electron (e_2). So, their electron–pair density would be a product of the individual electron densities.

Electrons, however, do carry charge and spin and that leads to a decrease in probability of finding e_1 near e_2. This reduced probability is called the *exchange–correlation hole*, which refers to the decreased probability of finding e_2 at a given position r_2 when e_1 is in the given position r_1. The exchange–hole, also called Fermi–hole, defines the static reduction of probability with respect to one electron. The correlation–functional, or the Coulomb–hole,

causes the probability of finding an electron to increase at further distances of reference, while it decreases for shorter distances of reference, given another reference electron.

Electron self–interaction is embedded in the Coulomb term or the correlational functional. In multi–electron physical systems, it is increasingly difficult to account for this self–interaction coming from the system's electrons explicitly. Improvements in DFT methods try to converge towards an exact cancellation of the self–interaction energy with the exchange energy. Currently, ongoing effort to develop the new and improved exchange–correlational functionals is trying to perfect that cancellation effect. To devise the absolute form of this exchange–correlational functional without breaking a functional is arduous. Also, theory shows that a local potential that is completely self–interaction free is non–existent.

Every density functional has a distinct exchange–correlational functional. Ultimately the system and its interesting properties dictate the appropriate functional we should use. Currently, the trend is to construct the E_{XC} as two distinct functionals for E_X and E_C, which when combined gives a physically meaningful exchange–correlation hole component. This initially separate construct of exchange and correlation functionals arises from the different scaling properties of the two energies.

As E_{XC} is an empirical functional, DFT methods do not have a systematic way to correct for the electron correlation. DFT improves upon the different E_{XC} functionals by *Jacob's Ladder*. Each functional in this ladder has improved accuracy in the following order: Local Density Approximation (LDA) < Generalized Gradient Approximation (GGA) < Meta-GGA < Hybrid– or Hyper–GGA < Double Hybrid methods and so on. Climbing a rung in this ladder means that the approximated functional is converging towards the exact form of the E_{XC} functional.

Correcting for dispersion (van der Waals forces) is important for accurate computation of molecular properties. An exact E_{XC} functional would exactly treat the dispersion interactions. As the approximated E_{XC} functional neglects this treatment, the *dispersion corrections* reintroduce the missing inter– and intra– molecular van der Walls forces into the functional. One of the major improvements in method development, in the mid–2000s, was the introduction of this dispersion correction inside the E_{XC} functionals.

3.0 Two Dimensional Infrared Spectroscopy Method

The 2D-IR spectroscopy experimental setup that generated the results in chapter 4 is described in this section. First we go through a short overview of the 800 nm light entering the optical setup to generate the mid-IR pulses. After that, the focus is on the white light generation, optical parametric amplification, and difference frequency mixing processes in an optical parametric amplifier (OPA). Finally, the general scheme of 2D-IR measurement process in the 2D spectrometer is presented, where the roles of the pulse generation, the motorized stages, the sample plane, and the detectors are elaborated.

A Ti:sapphire oscillator (Coherent Vitesse) pumps the amplifier at 80 MHz with 100 fs pulses ($\sim$ 4 nJ/pulse, Figure 3.1). The Ti:sapphire amplifier (Coherent Legend Elite) outputs 120 fs ultrashort pulses $\lambda = 803$ nm and BW = 13 nm, at a repetition rate of 5 kHz and an average power of 5 W. One watt of this amplified light is sent to the OPA.

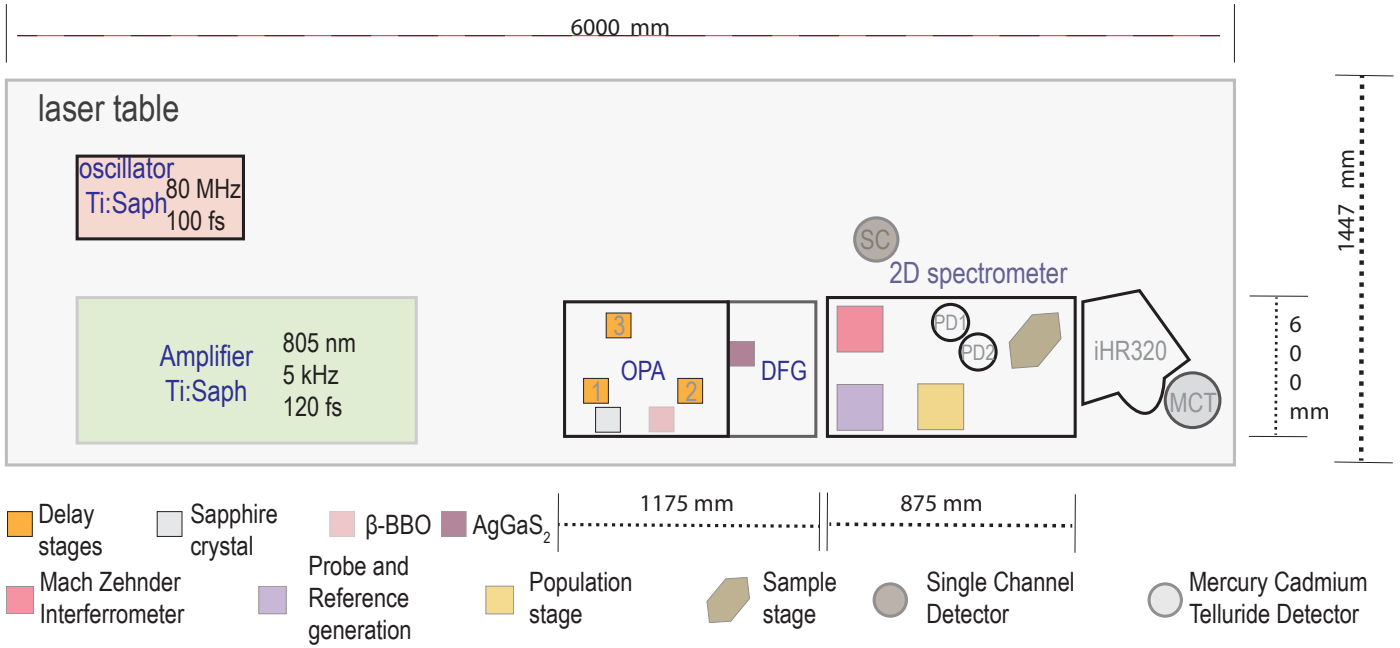

Figure 3.1: Optical setup for a 2D-IR measurement. The diagram shows the overall equipment, their dimension, and the text in this chapter goes into the relevant details.

3.1 Mid Infrared Light Generation

In the OPA, the 805 nm light goes through three paths. First is the white light generation, along with two pump pulses (Figure 3.2). The first section of the OPA *pre-amplifies* the desired signal frequency with the white-light seed. After passing once through the BBO crystal, the signal and the idler are generated from the pump pulse. The signal is taken and pumped again with higher energy during the second pass through the BBO. This step is called the *amplification* stage. After this pass, the optical parametric amplification process is complete. Finally, both the signal and idler go through the $AgGaS_2$ (AGS) crystal, which generates a new wavelength from the signal and idler by *difference frequency generation* (DFG). This new frequency is in the desired mid-IR range required for a particular 2D-IR experiment. Spatial and temporal overlaps of the pulses in each of the three processes are controlled by three delay stages.

Some materials have nonlinear optical responses to very high power ultrashort laser pulses. When the laser induced electric field becomes comparable to the material's electric field, we observe *white light continuum generation*. Sapphire, a centrosymmetric material with inversion symmetry, has a nonlinear index of refraction which is dependent on the laser intensity. During laser light exposure, *self-focusing* inside the crystal makes the beam center have a higher index of refraction than the rest of the crystal.

In our setup, one filament undergoes self-focusing and self-phase modulation, hence we have a *single-filament white-light continuum*. The white light frequencies have an inverse time dependence with the laser intensity. Our 805 nm input laser light source is nearly monochromatic. This time dependence creates a range of frequencies and colors at different spatial regions of the emergent beam. Therefore, the center of the white light has lower frequencies, while the edges have the high frequencies. The polarization of this generated white-light continuum is the same as the input pulse, horizontal.

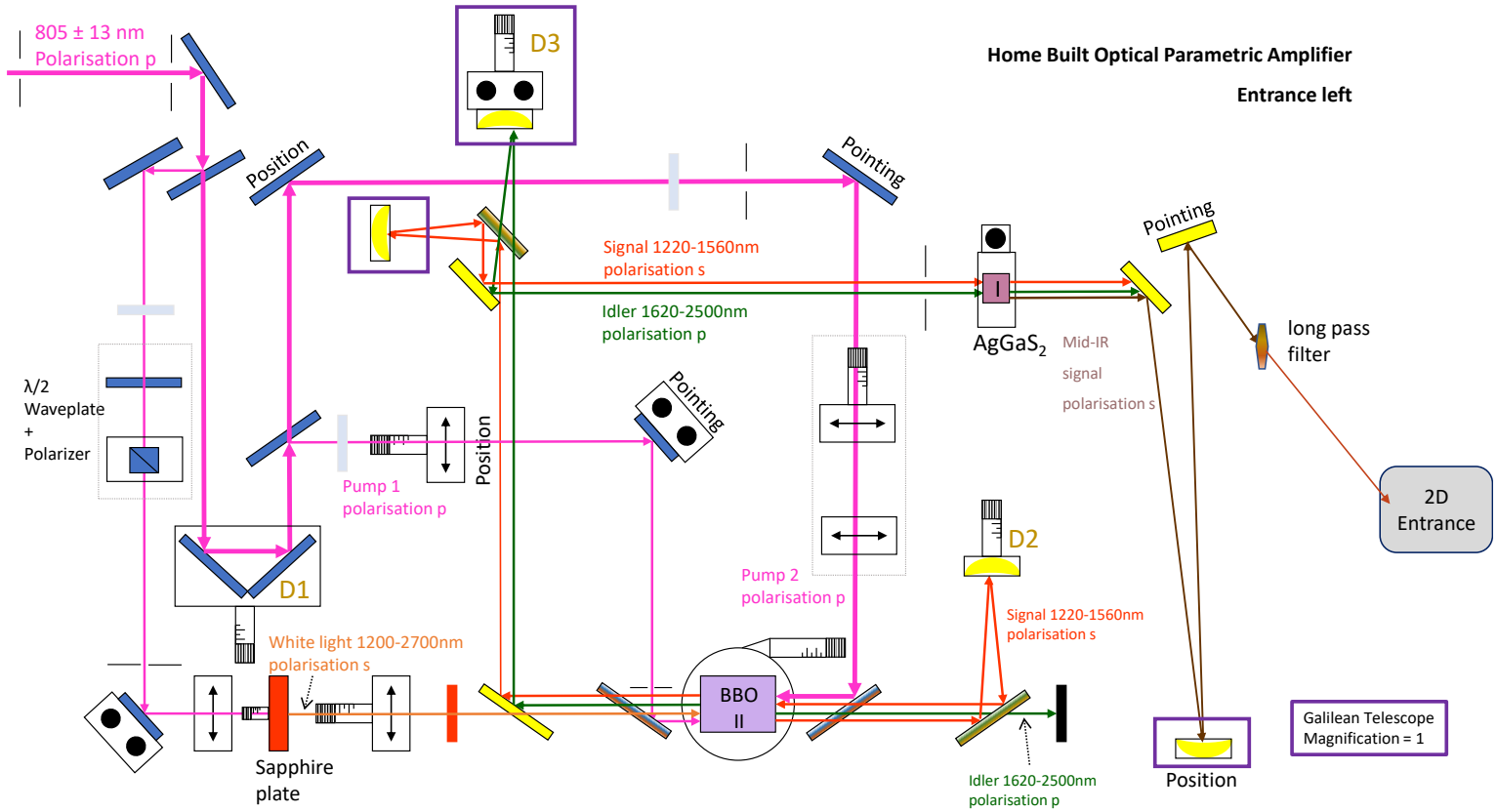

Figure 3.2: The optical parametric amplifier is based on the design by Hamm *et al.*[132,133] and built by Ren *et al.*[134]. The white light (light orange) from the sapphire crystal meets the pumps (pink) at the BBO to give rise to the signal (dark orange) and idler (green) pulses. After the second pass in the BBO, the signal and idler go through the AGS to generate the mid-IR pulse (light brown). The mid-IR pulse is filtered to eliminate all the NIR light before entering the 2D spectrometer. The beam splitters and reflective mirrors are colored in blue, the gold plated optics in yellow, dichroic mirrors in gradient color, and the delay stages are D1, D2, and D3.

A type-II β-barium borate (BBO) crystal is a negative uniaxial crystal. It is a colorless, trigonal, non-centrosymmetric second order material with the property of birefringence. As in white light generation, the electric field of the laser induces an electric field in the BBO crystal. A birefringent crystal has two axes, the ordinary and the extraordinary with indices of refraction η_{o} and η_{e} respectively. In the lab frame, the ordinary axis is oriented horizontally (H, parallel to the laser table) and the extraordinary axis vertically (V, perpendicular to the laser table). At a *phase–matching angle*, rays polarized along the ordinary and extraordinary axes have the same phase. At this phase–matching angle, rays polarized along the ordinary and extraordinary axes experience the same index of refraction.

The first pass through the BBO involves a white light seed and an 805 nm pump pulse. The white light is polarized in the direction of the BBO's extraordinary axis and the pump in the direction of the ordinary axis. With rotation stages, one can rotate the crystal about the extraordinary axis. This allows the projection of the ordinary axis onto the optic axis to be modulated. Thus, we are matching η_{e} along the extraordinary axis with η_{o}. The BBO crystal used in our setup decreases the beam velocity along the η_{e} axis and increases the beam velocity along the η_{o} axis. White light, aligned to the η_{e}, is phase matched to the pump aligned to the η_{o} axis of the BBO crystal. This is our desired angle for the generation of signal and idler from a white light seed. As the beam waist of the pump beam is small compared to the beam waist of the seed pulse, the latter gets spatially filtered giving only signal and idler outputs from the optical parametric amplification process. For example, in the 1400 cm^{-1} frequency region, the BBO phase–matching angle is found at a rotation stage setting of 331°. During the second pass through the BBO the pump pulse is still the 805 nm light, while the seed is the signal generated through first pass.

From the BBO, the signal and idler become the pump (η_{e}) and the signal (η_{o}) entering the $AgGaS_2$ crystal, respectively. This Type I $AgGaS_2$ crystal is a second order nonlinear non-centrosymmetric crystal performing difference frequency mixing. It is mounted on a vertical rotation stage allowing for rotation about the ordinary axis. The extraordinary axis is in the vertical (V) plane. For the Type I crystal, the output signal and idler photons are both aligned to the ordinary axis (as opposed to the extraordinary axis in Type II crystal).

When the phase–matching angle is achieved, the $AgGaS_2$ output is in the mid-IR region.

After the $AgGaS_2$, the new signal beam and the difference frequency output have the same polarization, η_o. This output mid-IR is the idler beam that enters the 2D spectrometer after NIR filtering. As an example, for a 1400 cm^{-1} frequency region experiment, the $AgGaS_2$ phase–matching angle is set at 183° in the rotation stage. Thin $AgGaS_2$ crystal plates (in our case 1.5 mm thickness) are popular for ultrashort pulse generation in the mid-IR range by DFG employing NIR wavelength pulses. Approximately 4% of the DFG is usable as our desired mid-IR wavelength range. In our setup, if 60 µJ energy enters the $AgGaS_2$ for DFG, then we get 2 µJ of output mid-IR into the 2D setup.

3.2 2D Spectrometer

From the OPA, the mid-infrared pulse enters the 2D spectrometer (Figure 3.3). After passing through a couple of mirrors, the beam diverges into three paths: the pumps, the probe, and the reference. The mid-IR beam's energy is in the range of 1–2.5 µJ/pulse, depending on the output frequency from the OPA. A Mach Zehnder interferometer generates the two pump pulses, while the probe and reference beams are created in a separate beam path. The pumps and the probe travel towards the sample plane to generate a 2D-IR signal. The 2D-IR and the reference signals are eventually collected by a mercury-cadmium-telluride (MCT) detector. Finally, the 2D-IR signal is gated, integrated, and digitized and transmitted to the lab computer by a Femtosecond Pulse Acquisition System, to be processed in MATLAB analysis software.

First, the mid-IR beam hits a CaF_2 wedge optic and encounters both reflective and transmissive surfaces in the optic's front and back. The transmitted light ($\sim$ 90% of the incoming energy) falls into a Mach Zehnder interferometer and generates the pumps. The Mach Zehnder interferometer has two arms: the *moving arm* and the *stationary arm* with two mirrors in each. The former sits on a motorized translation stage, while the latter is immobile, as the name suggests. In between the arms are two 50–50 beam splitters that guide the beam to reach both arms. Traveling through the beam splitters twice generates four pulses, two of each oriented at 90° and 180° to the incoming pulse. IR pump light into

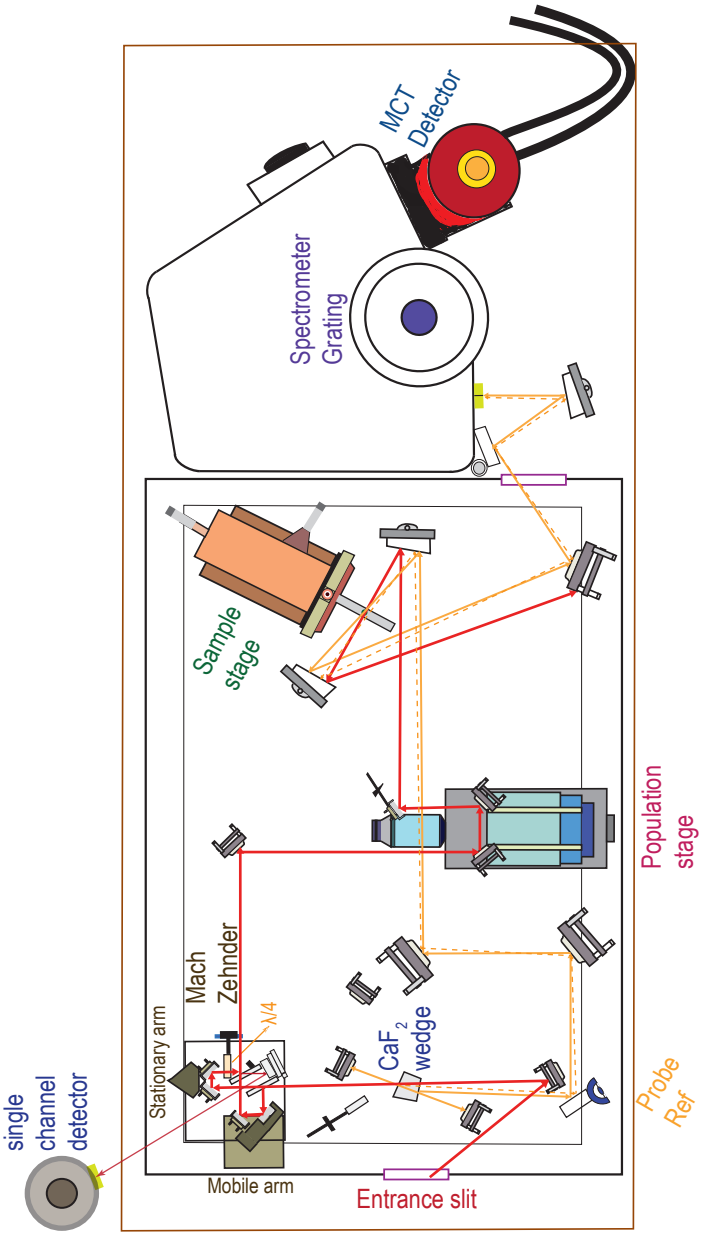

Figure 3.3: In the 2D spectrometer, the optical paths of the two pumps are indicated in red, while that of the probe is yellow and reference is dashed yellow. The ultimate destination of the 2D-IR signal and reference is the MCT detector, while a portion of the pump interference signal goes into the single channel detector.

the sample is from the two pulses that are turned 90° from the input mid-IR light in the interferometer. The total light intensity in both the 90° and 180° directions is $\sim 50\%$ of the input IR beam. As each of these two directions are separate for the two pumps, each pump has $50\% \times 50\% = 25\%$ energy of the original mid-IR beam.

After the interferometer, two ultrafast pump pulses with a t_1 time delay between them travel towards the population stage. A Fourier transform of this fast scanning t_1 axis generates the ω_1 frequency axis in our experiment. The second output of the Mach Zehnder interferometer is sent to a single-channel MCT detector. The single channel collects the interference pattern between the pumps that establishes the phasing of the 2D spectrum.

The Mach Zehnder interferometer contains another stack of mirrors on top of the stationary arm and the beam splitters. As the moving arm mirror is large in size it accommodates the HeNe beam on the top of that same mirror. The HeNe beam travels 2 cm^{-1} above the pumps. A $\lambda/4$ waveplate, which sits in front of the stationary arm in the HeNe path, converts the static arm's linearly polarized light to circularly polarized light. Thus the fixed arm creates a circularly polarized HeNe, and the moving arm HeNe is linearly polarized. After the interferometer, the HeNe beams fall into a 50–50 polarization beam splitter cube (PBSC), oriented at 45-degree angle to the incident polarization. Two photodiodes are placed orthogonally on two faces of the PBSC. When the circular and linearly polarized HeNe interfere, they produce a phase difference of 90° with each other. The two photodiodes collect these 90° phase-shifted HeNe signals.

We count the HeNe interferogram fringes (bins) to establish an equidistant time grid and obtain an unambiguous t_1 delay and pump phase difference. HeNe emission (632.8 nm) creates time compartments with a width of 2.11 fs called bins. Bin numbers sort the mid-IR data at the instant of laser pulse emission. On an average, 13–14 shots fall in one bin in our setup. Signals falling into a single bin are averaged. Convoluting a square window function with the time-domain signal is an analogue of this binning operation. The single-channel detector captures the interferogram between the two pump pulses. The plot of the single-channel interferogram and the bin number reflects the bins with the maximum pump intensity.

A CaF_2 wedged optic creates the probe and reference beams. Fresnel reflections off the front and rear surfaces of the CaF_2 send 5% of the mid-IR total energy to the top stacked

mirror (probe) from the front surface and 3% from the back surface (reference). Both the probe and reference beams travel in an almost parallel path only subtending a minor acute angle. This results in the probe and reference falling at vertically separated spots in the off-axis parabolic mirror (OAP) mirror before the sample stage.

The population stage has two mirrors on a mechanically controlled translation stage (t_2 delay). An OAP focuses the pumps and the probe into the sample, generating the third order nonlinear signal. The reference falls 100 µm above the pump and probe spots, in the sample plane. After the sample plane, the pumps are discarded while the probe and reference keep going to the spectrometer. Outside the 2D spectrometer box, two stacked mirrors receive the probe (bottom) and reference (top) beams. It then sends the probe and reference to an OAP. The focus of this OAP is right after the $\sim$ 2 mm wide vertical slit, inside the spectrograph. From there, the two beams travel towards the collimating mirror at one end of the spectrograph.

In our iHR320 spectrograph, the collimating mirror inside collimates the probe and reference onto the gratings. There are three grating options in our setup, 50 l/mm, 75 l/mm, and 150 l/mm. During the experiment, we choose one grating to set the frequency window and the ω_3 resolution of our experiment. From the diffraction grating turret, the beams fall on another focusing mirror, which sends it to the array detector. The IR signal is detected by a liquid N_2-cooled 2 $\times$ 32 mercury-cadmium-telluride (MCT) array detector. The probe falls on the top 1–32 elements, while the bottom 32–64 elements get the reference pulse. The grating produces the ω_3 final frequency axis of the 2D-IR spectra.

4.0 CH Mode Mixing Determines the Bandshape of Carboxylate Symmetric Stretch in Apo, Ca^{2+}, and Mg^{2+}–EDTA

This text is an adaptation from Mitra, S.; Werling, K.; Berquist, E.J.; Lambrecht, D.S.; and Garrett-Roe, S.; *J. Phys. Chem. A* **2021**, *125, 22*, 4867–4881. The author has contributed to this work by choosing the system of study, designing and conducting experiments, analyzing all experimental data, interpreting all the participation coefficient plots, setting up the initial DFT calculations and writing much of the manuscript.

4.1 Chapter Summary

The infrared spectra of EDTA complexed with Ca^{2+} and Mg^{2+} contain, to date, unidentified vibrational bands.This study assigns the peaks in the linear and two-dimensional infrared spectra of EDTA, with and without either Ca^{2+} or Mg^{2+} ions. Two-dimensional infrared spectroscopy and DFT calculations reveal that in both the presence and absence of ions, the carboxylate symmetric stretch and the terminal CH bending vibrations mix. We introduce a method to calculate participation coefficients that quantify the contribution of the carboxylate symmetric stretch, CH wag, CH twist, and CH scissor in the 1400–1550 cm^{-1} region. With the help of participation coefficients, we assign the 1400–1430 cm^{-1} region to the carboxylate symmetric stretch which can mix with CH modes. We assign the 1000–1380 cm^{-1} region to CH twist modes, the 1380–1430 cm^{-1} region to wag modes, and the 1420–1650 cm^{-1} region to scissor modes. The difference in binding geometry between the carboxylate-Ca^{2+} and carboxylate-Mg^{2+} complex manifests as new diagonal and cross-peaks between the mixed modes in the two complexes. The small Mg^{2+} ion binds EDTA tighter than the Ca^{2+} ion, which causes a redshift of the COOs modes of the sagittal carboxylates. Energy decomposition analysis further characterizes the importance of electrostatics and deformation energy in the bound complexes.

4.2 Introduction

Ethylenediaminetetraacetic acid (EDTA) can model divalent ion–carboxylate interactions in proteins.[52,135,136] In the broad class of EF-hand proteins, for example, divalent ion-binding induces structural changes[50] and begins a cascade of protein folding[27], cell signaling[27], ion transport[27], ion recognition[27], and enzyme activation events[137–141]. Binding selectivity and affinity of Ca^{2+} over Mg^{2+} is critical for the biological function of these proteins.

Though EDTA mimics the EF-hand binding pocket, spectroscopic studies fail to unambiguously correlate metal binding geometries with the positions and intensities of the carboxylate vibrational bands[142–148]. The additional information content of two dimensional infrared (2D-IR) spectroscopy promises to allow more direct relationships between observed spectral features and ion–binding geometries. Additionally, it can provide insight into the femtosecond and picosecond dynamics in the binding pocket.[46,94,95,149]

In the absence of divalent ions, EDTA has an extended, flexible structure in solution. Upon metal binding, EDTA assumes an approximately C_2 symmetry (Figure 4.1). The pairs of carboxylates lie in two planes. In the 'sagittal' plane (along the violet line), the two carboxylates are oriented parallel to each other, while the other two carboxylates are oriented antiparallel to one another and lie rotated slightly out of the 'equatorial' plane. The binding geometries of calcium- and magnesium-bound EDTA, $[Ca{:}EDTA]^{2-}$ and $[Mg{:}EDTA]^{2-}$, respectively, differ subtly. The smaller Mg^{2+} ion lies deeper in the binding pocket and the terminal oxygens approach the Mg^{2+} more closely. These differences cause changes in the NCCO dihedral angles along each of the carboxylates. These subtle rearrangements are similar to the structural differences between Ca^{2+}- and Mg^{2+}-bound EF-hands. If spectroscopic observables could be related to these differences in geometry, they might prove a powerful tool for determining the mechanism of ion selectivity in EF-hand proteins.

2D-IR spectroscopy is highly sensitive to local picosecond fluctuations, environmental dynamics, and molecular structure.[46,150] Techniques like X-ray crystallography[151] and EPR[58] and NMR[152] spectroscopies report the static structure and slower millisecond dynamics of ion binding[153–155]. The sub-picosecond temporal resolution of 2D-IR spectroscopy, however, can augment our understanding of the ultrafast dynamics in peptide–ion interac-

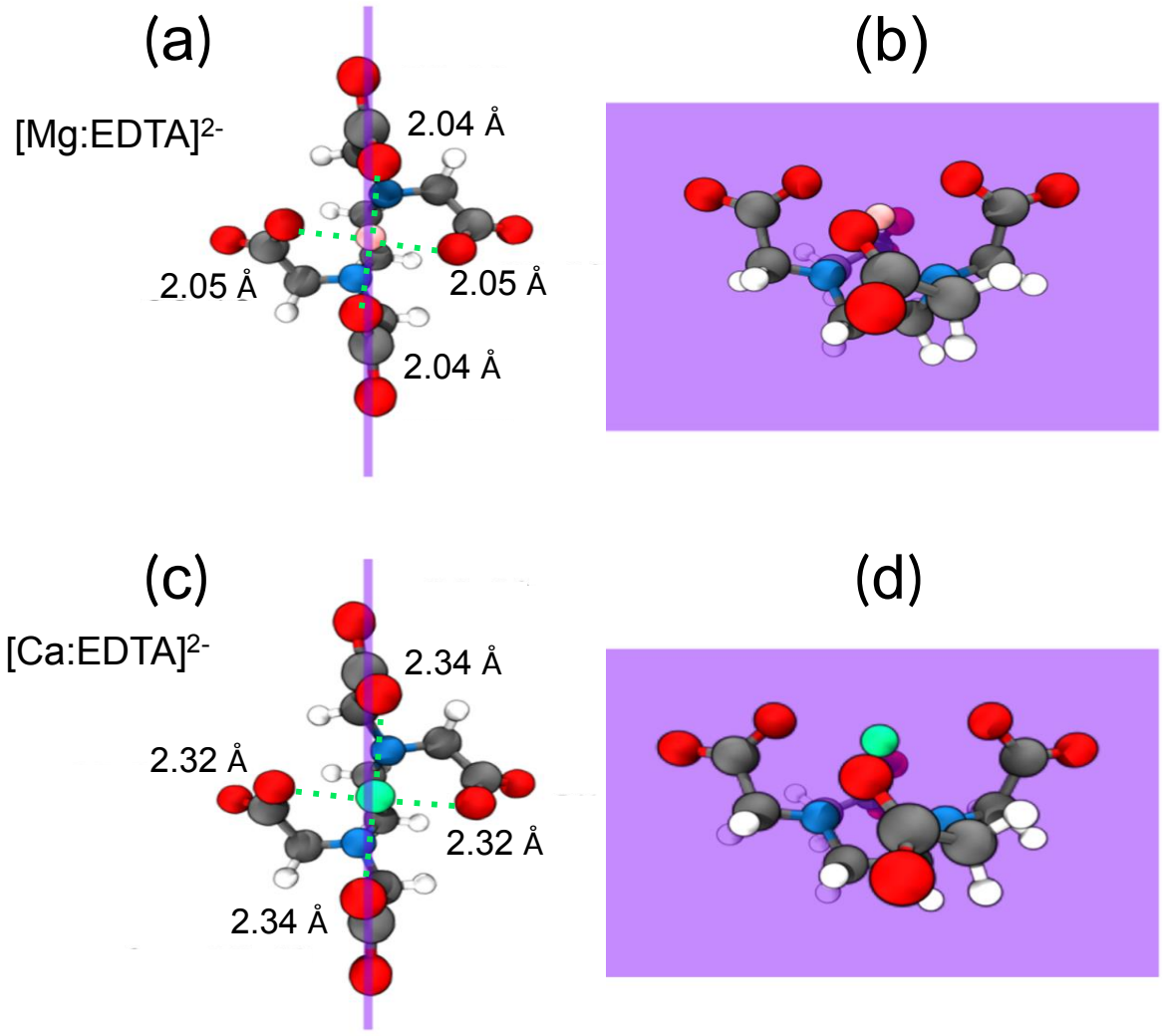

Figure 4.1: EDTA adopts an approximately C_2 symmetry when it binds Mg^{2+} (a, b) and Ca^{2+} (c, d). O–M^{2+} bond distances indicate a tighter binding geometry for $[Mg:EDTA]^{2-}$ than $[Ca:EDTA]^{2-}$. The sagittal acetates lie parallel to each other in a vertical plane (purple), while the other acetate pairs lie antiparallel to each other near the equator of the molecule.

tions.[36,45,46,78,90,156,157]

Infrared spectroscopy of EDTA probes the structure of carboxylate-Ca^{2+} ($[Ca{:}EDTA]^{2-}$) and carboxylate-Mg^{2+} ($[Mg{:}EDTA]^{2-}$) complexes[94,142,145,147,148]. Deacon and Phillips[158] empirically correlated the wavenumber difference between the antisymmetric stretch (COO^-_a) and symmetric stretch (COO^-_s) ($\Delta\nu_{a-s} = \nu_{COO^-_a} - \nu_{COO^-_s}$) with the binding geometry for aqueous metal acetato-carboxylate complexes. The $\Delta\nu_{a-s}$ roughly correlates to unidentate, bidentate, bridging, and pseudo-bridging binding types[147,159,160]. Later, however, Deacon and Phillips[161] noted the inadequacy of using the $\Delta\nu_{a-s}$ empirical relation to explain ion binding structures in multi-carboxylates. Complex multi-carboxylates' COO^-_s region exhibits additional features after metal binding[46,52,54,94,95,135,136,147,159,162–164], whose origins were unexplained.

2D-IR spectroscopy of the COO^-_a promises insights into the local structure and dynamics of carboxylate groups. Because ions and molecules in the local solvation shell modulate the vibrational frequencies of carboxylate stretches, 2D-IR spectroscopy can report the picosecond structural dynamics. Metal coordination–induced geometry changes are reflected in the vibrational couplings. When a metal binds, the vibrations couple to generate cross-peaks in the 2D-IR spectrum. Coupling cross-peaks in EDTA and calmodulin (CaM) COO^-_a uncover structural changes.[94,149] The vibrations couple, through space (electrostatically) or mechanically (covalently), due to the metal binding event. Binding–induced structural changes have also been reported in the 2D-IR spectra of the amide modes in ionophores such as valinomycin[85] and ion channels.[78,165,166] For example, the side-on and end-on ion–pair conformers of calcium oxalate interconvert on a picosecond timescale[167]. A model Hamiltonian for the COO^-_a was able to relate the conformational changes due to oxalate–metal interaction to the spectral features of the COO^-_a vibration.[83,84,167] In trifluoroacetate, solvent fluctuations drive population relaxation from the pumped COO^-_a into the COO^-_s.[83]

Finally, the ion-coordination geometry and fluctuations in the binding pocket of wild and mutated CaM were elucidated with 2D-IR spectroscopy.[94,95] Edington *et al.*[94,149] show that the COO^-_a reliably captures the ion–induced distortion in the binding geometry of EDTA and CaM. A comparison of Ca^{2+} and Ln^{3+} shows that the bound ion dictates if the CaM binding pocket will be compact and rigid or loose and flexible, which translates into downstream tertiary structural changes.[94] Separate active site mutations in CaM show

different antisymmetric stretch cross-peak structures, which are the manifestations of the CaM binding pocket's conformational flexibility.[95] In each of these examples, the COO^-_a bands in 2D-IR spectra provided important insights into local structure and dynamics of these carboxylate moieties.

While investigation of the COO^-_a reveals some changes upon ion binding[46,83,94,95,149], the symmetric stretch remains largely unexplored. Attenuated total reflection Fourier transform infrared (FTIR) spectroscopy shows that the COO^-_s can act as an indicator of transition metal ion binding in a self-assembled bilayer.[168] Donaldson *et al.*[169] probed the COO^-_s mode of a carboxylate-capped layer on gold nanoparticles to understand the correlation between the peptide-binding geometry and carboxylate orientation. Faint cross-peaks upon gold nanoparticle binding reflect the orientation of the carboxylates on the nanoparticle surface.

In part, the sparsity of studies using the symmetric stretch are due to the complex lineshape. The symmetric stretching band is complicated because the absorption frequencies of the COO^-_s mode and CH bending modes are similar. Absorption bands from the CH bending modes overlap with the COO^-_s band in EDTA in the absence of bound ions (apo-EDTA), making it challenging to separate their contributions.[149] As a result, the interpretation of features in the 1300–1500 cm^{-1} region is unclear, independent of metal ion binding.

The complexity of the COO^-_s band is also an opportunity. The substructures in the lineshape may reveal subtle conformational changes that the COO^-_a does not. We aim to connect the origin of the COO^-_s region vibrational structures to the metal–binding geometry of EDTA. We also explore the feasibility of utilizing the COO^-_s mode as a marker band for conformational changes in peptides.

We address the link between EDTA's COO^-_s IR spectral features and its binding conformation. Combining 2D-IR spectroscopy and density functional theory (DFT), we characterize the relationship between the vibrational spectrum of apo- and metal-bound EDTA and the geometry of binding. 2D-IR spectroscopy and DFT calculations show the impact of Ca^{2+} and Mg^{2+} binding in the COO^-_s region. To separate the contributions of the COO^-_s and CH bending modes in the 1000–1650 cm^{-1} region, we develop a scheme to calculate participation coefficients from the vibrational normal mode calculations. We use these participation coefficients to quantify the symmetric stretch and CH bending modes' local contribution to each

normal mode and, from this information, assign the EDTA bands between 1000–1650 cm^{-1}. Furthermore, we apply energy decomposition analysis (EDA) methods to distinguish the different interaction energy terms that play a significant role in EDTA–ion binding. We use EDA to decipher the physically meaningful intermolecular interactions that initiate and drive the observed EDTA geometry changes upon metal binding.

This paper is organized as follows: First, we show the changes in the FTIR and 2D-IR spectra of EDTA as a function of the divalent ion identity (section 4.4.1). Next, to untangle the CH bending modes and COO_s^- contributions in the symmetric stretch region, we develop a participation coefficient analysis from DFT normal mode calculations. The participation coefficients help us assign the COO_s^- peaks (section 4.4.2). With those assignments, we rationalize the major differences in the 2D-IR spectra between the apo-EDTA, $[Ca:EDTA]^{2-}$, and $[Mg:EDTA]^{2-}$ COO_s^- and link them to the binding geometry of each ion (section 4.4.3). Finally, we use EDA calculations to fully characterize the nature of the intermolecular interactions that drive the transformation into the metal–bound conformation in EDTA (section 4.4.4).

4.3 Methods

4.3.1 Materials

Ultra pure grade $\geq$ 99.5% tetra-sodium salt of EDTA ($C_{10}H_{12}N_2O_8Na_4{\cdot}H_2O$) was purchased from AMRESCO, Inc. Calcium chloride ($CaCl_2$) and magnesium chloride ($MgCl_2$) salts were obtained from Sigma-Aldrich, Inc. Deuterium oxide (D_2O) solution was bought from Cambridge Isotope Laboratories, Inc. Deuteration eliminates the overlap between the carboxylate stretch region and H_2O bending mode. To avoid metal contaminants from the glass surface containers via leaching[170], EDTA solutions were made in 15 mL Falcon tubes. The powder form of EDTA and respective metal ions were directly mixed in D_2O to obtain the desired concentration. The uncorrected pH reading of the tetra-sodium EDTA salt was ~11.2. FTIR spectroscopy (Figure 4.2) verified the complete deprotonation of the EDTA

carboxylates in the aforementioned pH.

The 2D-IR samples were (1) 350 mM $[\text{apo-EDTA}]^{4-}$ in D_2O (2) 450 mM $[\text{Mg:EDTA}]^{2-}$ with 450 mM $MgCl_2$ in D_2O and (3) 460 mM $[\text{Ca:EDTA}]^{2-}$ with 460 mM $CaCl_2$ in D_2O solutions. These concentrations provided absorption in the 0.20–0.40 OD range for 2D-IR measurements.[29] Each FTIR and 2D-IR sample cell had a 15 µL solution droplet in between two 2 mm thick CaF_2 UV-grade windows separated by a 6 µm polytetrafluoroethylene spacer.

4.3.2 Linear IR

A Thermo Fisher Nicolet 6700 spectrometer recorded each sample's FTIR spectrum with $2\,\text{cm}^{-1}$ spectral resolution. The sample chamber was purged with nitrogen during data collection to eliminate water vapor.

4.3.3 2D-IR

A commercial, 5 kHz Ti:Sapphire laser (Coherent Legend Elite) generates ∼120 fs pulses of 805 nm light with $\sim 250\,\text{cm}^{-1}$ full-width at half-maximum (FWHM) bandwidth and 1 mJ/pulse. An optical parametric amplifier generates mid-infrared pulses[133]; the amplified signal and idler pulses are mixed in a difference frequency mixing crystal, $AgGaS_2$, to generate $1.80\,\mu\text{J/pulse}$ at $1410\,\text{cm}^{-1}$.

The pump-probe geometry 2D-IR spectrometer is based on the design of Helbing *et al.*[171] The spectrometer generates two collinear pumps, a probe, and a reference pulse. A fast-scanning Mach-Zehnder interferometer generates the first coherence time, t_1, between the pumps. The population stage produces the waiting times, t_2, between the pump and the probe pulses. A Fourier transformation of the t_1 time at each t_2 gives the initial frequency axis, (ω_1). The signal field, which is self-heterodyned with the probe pulse, was diffracted by a 50 l/mm grating to disperse the ω_3 final frequencies.

A 2×32 mercury cadmium telluride detector (Infrared Associates, liquid N_2 cooled) collects the signal and reference. Signals are gated, integrated, digitized, and transferred to the computer using a Femtosecond Pulse Acquisition Spectrometer (Infrared Systems Development Corps.). MATLAB (MathWorks) is used to visualize and analyze the collected

2D-IR signal. In the current work, all the 2D-IR spectra are at $t_2 = 200\,\text{fs}$ with a signal averaging of ∼600 scans each and with resolutions of $6\,\text{cm}^{-1}$ along ω_1 and $3\,\text{cm}^{-1}$ along ω_3.

4.3.4 Calculations

4.3.4.1 Density functional theory

Geometry optimization and vibrational frequency calculations were performed on apo-EDTA (32 atoms) and $[\text{Ca:EDTA}]^{2-}$ and $[\text{Mg:EDTA}]^{2-}$ (33 atoms) with the M05-2X[172] and ωB97X-D[173,174] density functionals using a development version of Q-Chem[175]. All calculations use the def2-TZVP basis set. Standard grids 3 and 2, as implemented in Q-Chem[176], were used for all M05-2X and ωB97X-D calculations respectively. The condensed phase environment was incorporated via the SMD[177] implicit polarizable continuum solvation model for all three molecules. The three molecules were generated in Avogadro[178] with an initial force-field based geometry optimization (MMFF94s with steepest descent), from which the resulting Cartesian coordinates were used as input to the DFT geometry optimizations. ChElPG-based atomic partial charges of the metal-bound Mg^{2+} and Ca^{2+} were performed on the optimized structures. Computed frequencies used in the paper are unscaled, and the scaling factors with their respective scaled frequencies were calculated for both functionals (Appendix A.1).

The M05-2X functional was chosen because it accurately represents carboxylate systems, especially for thermodynamic information.[179,180] The remaining differences between computed and experimental spectra generally originate from incomplete treatment of the condensed phase environment[181,182] and anharmonicity. More specifically, gas-phase single-conformer calculations are missing solute-solvent interactions and contributions from energetically low-lying conformers. We approximate the solute-solvent interactions using an implicit solvent model in order to avoid the costly sampling required for incorporating explicit solvation. Currently, a test set of small molecules, such as acetate and citrate, are used to benchmark carboxylate spectra at the comparable M05-2X/cc-pVTZ/SMD level of theory[180], but these small benchmarks may still not consider the complicated interactions of molecules as large as EDTA; for example, the M05-2X functional is reported to overestimate the COO^- symmetric

stretch frequencies compared to experiment. As comparison, we supplement the use of the M05-2X functional with ωB97X-D, since previous work[183,184] shows similarities between the two functionals for harmonic frequencies.

4.3.4.2 Intermolecular interactions

To further understand the nature of the metal-ligand binding, we quantified this intermolecular interaction using EDA based on absolutely localized molecular orbitals (ALMO-EDA), which quantifies the interaction based on physically intuitive components. In ALMO-EDA with implicit solvent, the total change in energy due to bringing isolated fragments together is given by

$$\Delta E_{\text{int}}^{(\text{s})} = \Delta E_{\text{geom}}^{(\text{s})} + \Delta E_{\text{frz}}^{(\text{s})} + \Delta E_{\text{pol}}^{(\text{s})} + \Delta E_{\text{CT}}^{(\text{s})} \tag{62}$$

$$= \Delta E_{\text{geom}}^{(\text{s})} + \left(\Delta E_{\text{elec}}^{(\text{s})} + \Delta E_{\text{Pauli}}^{(\text{s})} + \Delta E_{\text{disp}}^{(\text{s})}\right) + \Delta E_{\text{pol}}^{(\text{s})} + \Delta E_{\text{CT}}^{(\text{s})}, \tag{63}$$

where ΔE_{geom} is the energy increase caused by the fragments adopting their interacting geometries, ΔE_{frz} is the energy change from bringing fragments together into the final supermolecular geometry while keeping their densities frozen and not allowing them to relax, ΔE_{pol} is the energy lowering from allowing intrafragment density relaxation in the presence of other fragments while disallowing interfragment charge flow, and ΔE_{CT} corresponds to the charge transfer caused by lifting this last restriction. The frozen density term can be further broken down into three primary parts: ΔE_{elec} is the electrostatic Coulomb interaction between isolated fragments, ΔE_{disp} is the dispersion interaction, and ΔE_{Pauli} is the remaining short-ranged non-electrostatic components of bringing fragments together while keeping their densities constrained. The superscript "(s)" signifies that a term includes solvent effects.

The ALMO-EDA formalism is based on the classical decomposition of the frozen density[185,186] with solvation[187] as implemented in Q-Chem. As validation, we also performed symmetry-adapted perturbation theory (SAPT) calculations, which originate from a different theoretical foundation but contains comparable terms. These calculations are based on the SAPT0 formulation[188,189] as implemented in Psi4[190]. All ALMO-EDA and SAPT calculations used the same geometries as the ωB97X-D/def2-TZVP/SMD frequency calculations.

ALMO-EDA uses this method chemistry for all fragment wavefunctions, including the solvent model. SAPT instead starts from a Hartree–Fock reference. Due to technical problems with including solvation, all SAPT calculations were performed in the gas phase. However, we see that trends in like terms between ALMO-EDA and SAPT are qualitatively similar (*vide infra*). To better compare against ALMO-EDA, monomer-basis SAPT results will be presented in the main text, and dimer-basis SAPT results are in Appendix A.1. In order to perform this comparison, we modify equation 62 slightly to

$$\Delta E_{\text{int}} = \Delta E_{\text{geom}}^{(\text{s})} + \Delta E_{\text{elec}}^{\text{cls}(0)} + \Delta E_{\text{solv}}^{\text{el}} + \Delta E_{\text{Pauli}}^{\text{mod(s)}} + \Delta E_{\text{disp}}^{(\text{s})} + \Delta E_{\text{pol}}^{(\text{s})} + \Delta E_{\text{CT}}^{(\text{s})} \tag{64}$$

$$= \Delta E_{\text{geom}} + \Delta E_{\text{elec}} + \Delta E_{\text{solv}} + \Delta E_{\text{Pauli}} + \Delta E_{\text{disp}} + \Delta E_{\text{pol}} + \Delta E_{\text{CT}}, \tag{65}$$

where the classical decomposition is applied and the solvent correction to the electrostatic interaction is considered separately (see equation 9 in ref. 187). The superscript "(0)" indicates the term is calculated without solvent effects.

4.3.4.3 Participation coefficients

Participation coefficients P were calculated to determine the contribution of CHH and COO motions to the normal modes of the apo- and complexed forms of EDTA. In general, P is the square of the overlap of a unit basis vector on the EDTA molecule (apo or complex) normal mode vector in the subspace of a set of atoms for the molecule. The participation coefficients can be represented generally by:

$$P_{i,j} = \frac{\left| \langle {}_{basis}\vec{\nu}_i | {}_{mol}\vec{\nu}_j \rangle \right|^2}{\langle {}_{basis}\vec{\nu}_i | {}_{basis}\vec{\nu}_i \rangle}. \tag{66}$$

${}_{basis}\vec{\nu}_i$ and ${}_{mol}\vec{\nu}_j$ refer to the i-th basis normal mode (we will discuss our notion of basis shortly) and the j-th molecule normal mode (determined from the frequency calculations described in the previous section), respectively. For our purposes, $P_{i,j}$ will always be calculated for a subspace of the full nuclear position coordinates of the molecule. The subspace of coordinates will always consist of the carboxylate COO atom positions or the terminal CHH atom positions of a specified acetate group on EDTA. Note that only the basis vector in the subspace of atoms is normalized, and hence the participation coefficient is bounded by the

square length of the normal mode in the subspace of the specified CHH or COO coordinates (eq. 67). If index l spans this subset of nuclear position coordinates, S, then we can express this bound as

$$P_{i,j} \leq \sum_{l \in S} {}_{mol}\nu_{jl}{}^2, \tag{67}$$

where j and i are defined as above.

We were particularly interested in calculating participation coefficients for the carboxylate regions ($S = \text{COO}$) and the alpha carbon regions ($S = \text{CHH}$) in the EDTA molecules since CHH motions tend to couple with COO symmetric stretch regions. The basis vectors for these two regions were created according to the following procedure. We desired a complete basis to fully describe the motion of these atoms as they contribute to the overall normal modes. To this end, we created orientation-dependent rotation and translation normal modes in the subspace of the CHH or COO atoms for each of the four acetates in each molecule. Since the equilibrium positions for the three atoms specify a plane, the normal to this plane specifies one of the basis vectors used to create the rotation and translation normal modes. The midpoint bisector (drawn from the carbon atom to the midpoint between the two hydrogen/oxygen atom positions) in the CHH/COO plane was used as the second basis vector, and the cross product of the the previous two basis vectors determined the third basis vector. Given the normalized basis vector $\vec{e}_i$, the corresponding normalized mass weighted translation vector (for a CHH group as an example), $\mathbf{T}_i$, is

$$\mathbf{T}_i = \frac{(\sqrt{m_C} \cdot \vec{e}_i, \sqrt{m_{H_1}} \cdot \vec{e}_i, \sqrt{m_{H_2}} \cdot \vec{e}_i)}{(\sqrt{m_C + m_{H_1} + m_{H_2}})}. \tag{68}$$

For the rotational vectors, the moment of inertia tensor, $\mathbf{I}$, was calculated for the CHH and COO atoms of each acetate using the positions relative to the center of mass of the three atom system, and the tensor was diagonalized to yield the three angular velocity vectors (eigenvectors), $\vec{\omega}_i$. These vectors were ordered by their maximal overlap with the three translational basis vectors $\vec{e}_i$ given above. The normalized mass weighted rotational vectors (for a CHH group as an example) are then given by

$$\mathbf{R}_i = \frac{(\sqrt{m_C} \cdot \vec{r}_C \times \vec{\omega}_i, \sqrt{m_{H_1}} \cdot \vec{r}_{H_1} \times \vec{\omega}_i, \sqrt{m_{H_2}} \cdot \vec{r}_{H_2} \times \vec{\omega}_i)}{\sqrt{(\omega_i)}}, \tag{69}$$

where ω_i is the corresponding eigenvalue to the eigenvector of the moment of inertia tensor. Note that the coordinates $\vec{r}$ here are given relative to the center of mass ("com") of the three atoms (*e.g.*, $\vec{r}_C = \vec{x}_C - \vec{r}_{\text{com}}$).

The vibrations were determined from the diagonalization of the mass weighted Hessian for the toy harmonic potential given in equations 70 and 71 and were ordered by eigenvalue:

$$E = k_{CH/CO} \sum_{i,j \subset S} (|\vec{r}_i - \vec{r}_j| - r_{ij}^0)^2 + k_{\angle CHH/COO}(\theta - \theta^0)^2 \tag{70}$$

where θ is defined as:

$$\theta = \arccos\left(\frac{(\vec{r}_{H_1} - \vec{r}_C) \cdot (\vec{r}_{H_2} - \vec{r}_C)}{|\vec{r}_{H_1} - \vec{r}_C||\vec{r}_{H_2} - \vec{r}_C|}\right). \tag{71}$$

$k_{CH/CO}$ is the force constant of the CH/CO bond, while $k_{\angle CHH/COO}$ is the force constant for the H-C-H/O-C-O angle. Since only the relative values for these force constants are important, DFT calculations on methane (see Appendix A.2) indicate that the spring constant for the C-H stretch is about 4 times that of the HCH bend force constant, and so $k_{CH/CO} = 4$ and $k_{\angle CHH/COO} = 1$ were chosen for the model potential. (The value $k_{CH/CO} = 4$ was determined crudely from the DFT calculations for methane from the frequencies, f and reduced masses m to determine $k_{CH/CO} = \frac{k_{stretch}}{k_{freq}}$ from $\frac{f_{stretch}}{f_{bend}} = \sqrt{\frac{k_{stretch} m_{bend}}{k_{bend} m_{stretch}}}$). $\vec{r}_i$ and $\vec{r}_j$ correspond to the position of the i^{th} and j^{th} atoms. r_{ij}^0 is the equilibrium bond distance between the pair of atoms i and j. S is the set of pairs of atoms that are bonded to one another. θ is calculated from equation 71 as the angle between $\text{H}_1 - \text{C} - \text{H}_2$, where $\vec{r}_{\text{H}_1}$, $\vec{r}_{\text{H}_2}$, and $\vec{r}_{\text{C}}$ are the positions of H_1, H_2 and C. Here r^0 and θ^0 are the equilibrium position and angles. The resulting 9 basis vectors form a complete and orthonormal basis in this subspace.

4.4 Results and Discussion

4.4.1 Carboxylate Stretch Infrared Peaks in Apo-EDTA Change Shape after Metal Binding

Carboxylates absorb in the 1300–1700 cm^{-1} region (Figure 4.2). In apo-EDTA, the COO_a^- absorbs twice as strongly as the COO_s^-. The COO_a^- (1585 cm^{-1}) has a broad, roughly Gaussian

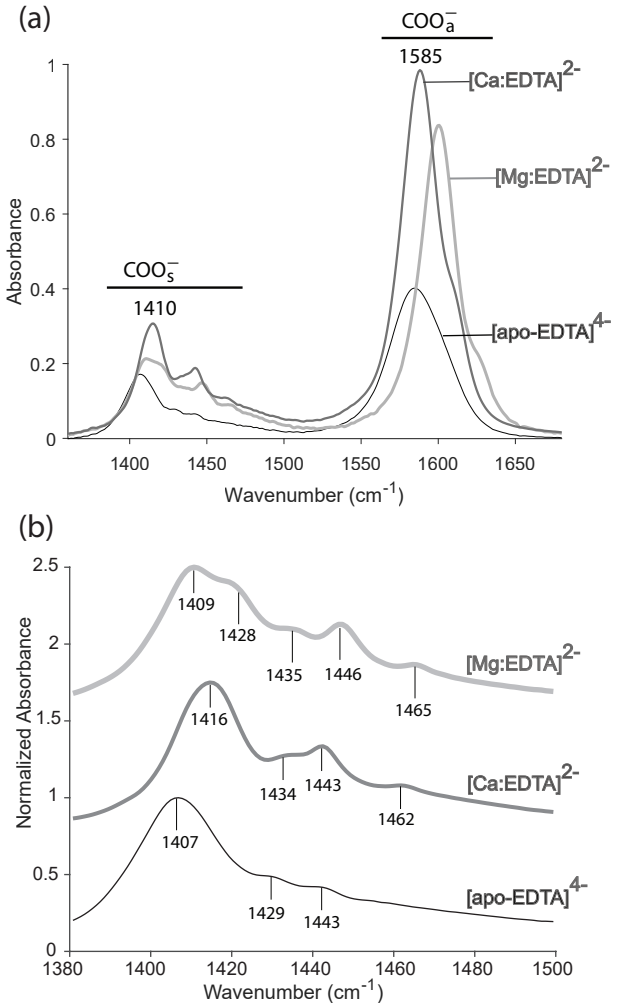

Figure 4.2: (a) FTIR spectra of apo-EDTA, $[Ca:EDTA]^{2-}$, and $[Mg:EDTA]^{2-}$ reveal the frequency, structure and lineshape changes in the COO^- stretch region. (b) The normalized spectra of the COO_s^- region have the peaks labeled for the three molecules (offsets: $[Mg:EDTA]^{2-}$ 1.5 A.U. and $[Ca:EDTA]^{2-}$ 0.75 A.U.)

peak shape (44 cm^{-1} FWHM), which agrees with literature[54,144]. The COO^-_s (1410 cm^{-1}) band, on the other hand, is broad (~42 cm^{-1} FWHM) and asymmetrical (absorption extends to nearly 1500 cm^{-1}). On top of the broad COO^-_s peak are lower intensity peaks at 1429 cm^{-1} and 1443 cm^{-1}, as previously observed.[144,145,159]

When EDTA binds Ca^{2+} or Mg^{2+}, the carboxylate stretches change their peak positions, structures, and lineshapes[144,145,159] (Figure 4.2). In the metal–bound EDTA, the COO^-_a band narrows to a Lorenzian ($[Ca:EDTA]^{2-}$: 31 cm^{-1} FWHM, $[Mg:EDTA]^{2-}$: 29 cm^{-1} FWHM), with a shoulder on the high frequency side.[149] The COO^-_s band blueshifts to 1416 cm^{-1} in $[Ca:EDTA]^{2-}$ and to 1409 cm^{-1} in $[Mg:EDTA]^{2-}$. In $[Ca:EDTA]^{2-}$, three new peaks grow at 1434 cm^{-1}, 1443 cm^{-1}, and 1462 cm^{-1}, while in $[Mg:EDTA]^{2-}$, four new peaks appear after complexation (1428 cm^{-1}, 1435 cm^{-1}, 1446 cm^{-1}, 1465 cm^{-1}).

2D-IR spectra add additional information about the coupling between the COO_s bands. At the earliest waiting time (t_2 = 200 fs), both diagonal and cross-peaks appear in the COO^-_s region of the 2D-IR spectrum of apo-EDTA (Figure 4.3 (a)). The diagonal peak $\omega_1 = \omega_3 = 1407\,cm^{-1}$ is inhomogeneously broadened, stretches along the diagonal, and has an elliptical shape. Faint diagonal peaks appear at $\omega_1 = 1429\,cm^{-1}$ and $\omega_1 = 1443\,cm^{-1}$. In addition to the diagonal band, a negative cross-peak, A, at $\omega_1 = 1407\,cm^{-1}$ stretches from ω_3 = 1423–1450 cm^{-1}, a negative cross-peak, B, ($\omega_1 = 1423\,cm^{-1}$) spans from ω_3 = 1407–1420 cm^{-1}, and a positive cross-peak, C, stretches from ω_3 = 1380–1400 cm^{-1}.

The 2D-IR spectrum of the COO^-_s vibrations changes upon metal binding to EDTA (Figure 4.3). At the earliest waiting time (t_2 = 200 fs), new diagonal bands and cross-peaks become apparent after metal binding. Consistent with the FTIR spectra, three diagonal bands in $[Ca:EDTA]^{2-}$ and four in $[Mg:EDTA]^{2-}$ are apparent. Additionally, the broad cross-peaks in apo-EDTA (A-C) resolve into sharp features in $[Ca:EDTA]^{2-}$ (A-D) and $[Mg:EDTA]^{2-}$ (A-F).

Our objective is to understand the origin of these vibrational bands in the FTIR and 2D-IR spectra after Ca^{2+} and Mg^{2+} ions bind to EDTA. In the following section, we will assign the vibrational bands utilizing the participation coefficients.

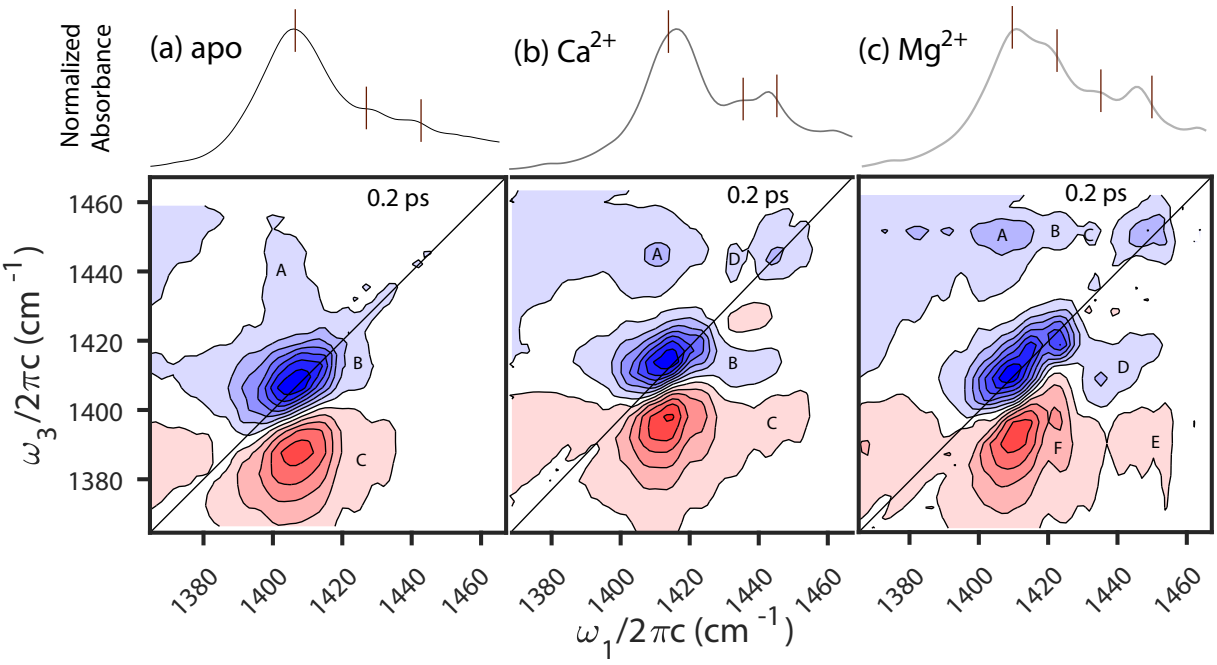

Figure 4.3: 2D-IR at $t_2 = 200\ fs$ spectra show the diagonal and cross-peaks in the three molecules' COO_s^- region: (a) 350 mM apo-EDTA (b) 460 mM $[Ca{:}EDTA]^{2-}$ (c) 450 mM $[Mg{:}EDTA]^{2-}$. For all three molecules, the bottom 2D-IR panel has cross-peak features labeled in the blue and red lobes. The top panel is the normalized FTIR spectrum of each molecule with peak centers aligned to the respective 2D-IR center frequencies.

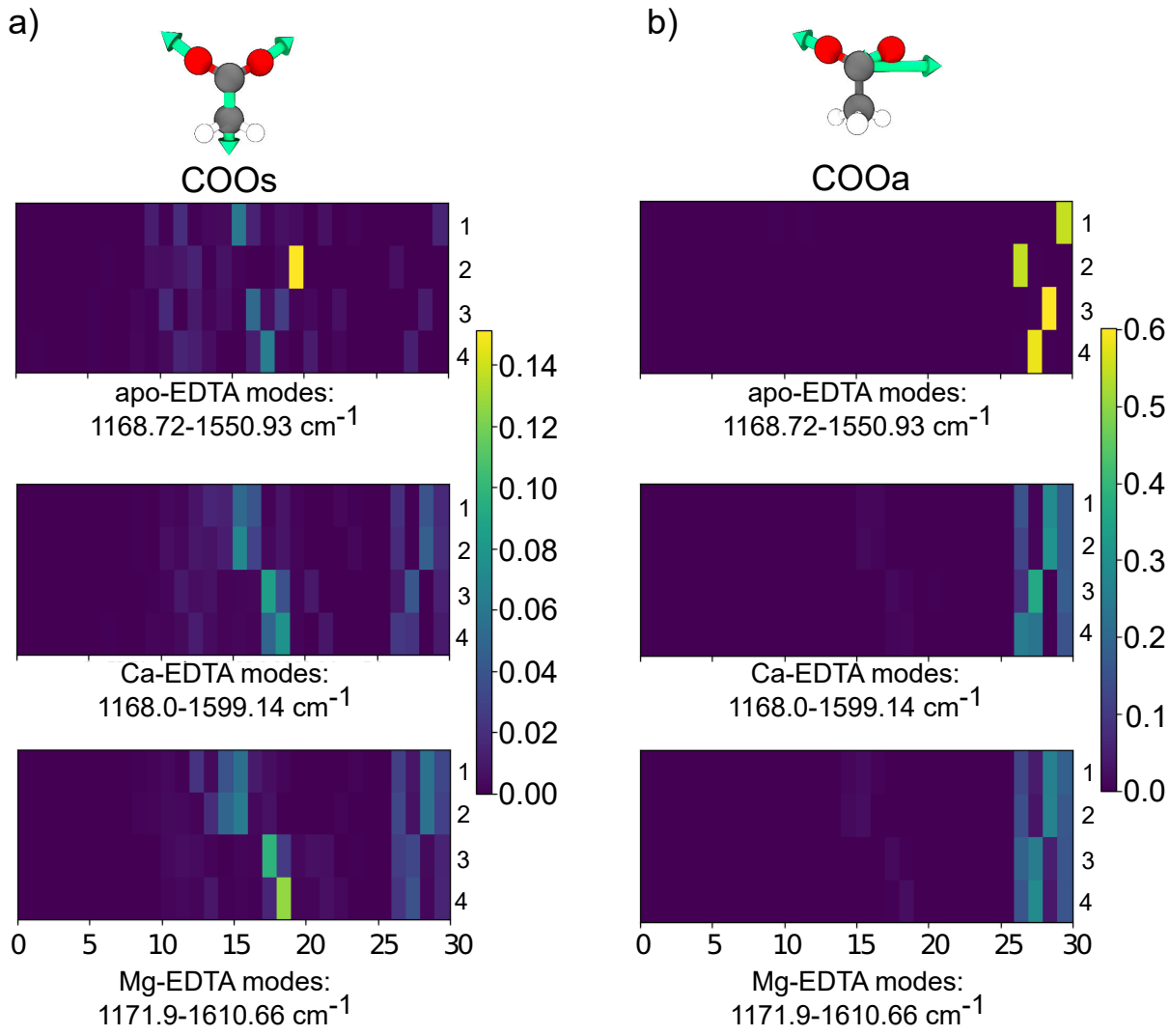

Figure 4.4: Participation coefficients for the COO_a^- are higher than the COO_s^-. Each column in the COO plot shows that COO_s^- is spread across more than four normal modes, unlike the COO_a^- which is limited to four modes. The top panel is apo-EDTA, $[Ca:EDTA]^{2-}$ is the mid-panel, and the bottom panel is $[Mg:EDTA]^{2-}$. The frequency range and indices are consistent throughout the paper. The left and right columns are COO_s^- and COO_a^- participation respectively. 1 and 2 are the sagittal acetates and the equatorial acetates are 3 and 4. All scales can be converted to a percent participation by multiplying with a factor of 100.

4.4.2 COO_s^-, CH_s and CH_w Mix in Apo-EDTA and Its Complexes

In this section, the entire 1000–1630 cm^{-1} range is assigned. We will first walk through the assignment of the COO_a^- band using the participation coefficient method, which will assist our understanding of the mixed COO_s^- assignments. Finally, we will discuss the nuances in the spectra of the three molecules. We will also address some of the unresolved assignments seen in the literature.

The calculated COO participation coefficients show the contribution of the idealized carboxylate vibrations to the normal modes of EDTA and the metal complexes for symmetric (Figure 4.4a) and antisymmetric (Figure 4.4b) stretches. In each panel, the x-axis indices are the normal modes from the respective three molecules (Table Appendix A.1) and the y-axis indices (1–4) are the four acetate groups in EDTA. In apo-EDTA (Figure 4.4b, top), only the normal modes 26–29 have appreciable projection onto the idealized COO_a^- modes. In addition, each of these normal modes projects only onto one of the localized COO_a^- stretches, which shows that the vibrations are localized in apo-EDTA. In $[Ca:EDTA]^{2-}$ (Figure 4.4b, middle), the COO_a^- participation coefficients also indicate that only normal modes 26–29 have appreciable COO_a^- character. The normal modes each project onto several of the idealized COO_a^- vibrations, however, indicating that the normal modes are delocalized across several carboxylates. In $[Mg:EDTA]^{2-}$ (Figure 4.4b, bottom), the same general patterns emerge, and modes 27 and 28 each are nearly equally delocalized across a pair of carboxylates. We will return to the structural origin of this observation later in this section. In summary, the participation coefficients show the projection of each normal mode onto an idealized basis vector, which can quantitatively illuminate the character of the normal mode and its degree of delocalization. In this case, we observe that only four normal modes have COO_a^- character, and metal binding can induce delocalization of the vibrations, as expected.[149]

Having introduced the relatively straightforward COO_a^- modes, we turn now to the COO_s^- modes, which are more complicated. In the COO_s^- of apo-EDTA (Figure 4.4a, top), more than four normal modes project onto the idealized COO_s^- vibration. The participation coefficients are appreciable from modes 10–19, indicating that the ideal symmetric stretch vibration is present in all these modes, while modes 15–19 overlap the most. In both

$[Ca:EDTA]^{2-}$ and $[Mg:EDTA]^{2-}$ (Figure 4.4a, middle and bottom), modes 14–18 overlap the most with the idealized COO_s^- vibration. Some contribution of modes > 25 is noted, which we attribute to the lower symmetry of each carboxylate upon ion binding, *i.e.*, symmetric and antisymmetric are no longer rigorous symmetry labels because of the nearby ion. Even with this complication, the participation coefficients allow us to identify normal modes 14–18 as the primary carboxylate symmetric stretches in EDTA and its metal complexes.

The COO_s^- motions are mostly localized, independent of metal binding (Figure 4.4a). The mixed COO_s^- are localized to each carboxylate. The highest participation from one COO_s^- in the 1464 cm^{-1} calculated band is ∼15%, and for the rest of the individual COO_s^- modes it is ∼7%. With Ca^{2+} binding, the highest COO_s^- participation redistributes from 7 to 9%, whereas in Mg^{2+} the individual COO_s^- participation falls in a broad range from 3–13%. The carboxylates, however, (Figure 4.4a) participate alone (modes 14–18), implying that only one carboxylate is heavily involved in the symmetric stretch vibration. Metal binding more or less does not change the localization of the COO_s^- vibration. We interpret that the insignificant delocalization in the metal complexes is from their low C_2 symmetry, while the disordered apo-EDTA structure produces localized motions. We will examine how the structure of the EDTA affects the 2D-IR spectra in-depth in the next section (Section 4.4.3).

Projecting the normal modes of apo- and complexed EDTA onto the idealized acetate CHH normal modes (twist (CH_t), wag (CH_w), and scissor (CH_s)) reveal their contribution in the 1000–1650 cm^{-1} region (Figure 4.5). In the CHH participation coefficients, two rows of information are available for each of the twist, wag, and scissor motions. For each of them, the top row, labeled Avg, includes the mean participation from all four CHH proximal to the carboxylates, whereas the four rows below the average show the participation of each of these four terminal CHH for CH_t (top panels), CH_w (middle panels), and CH_s (bottom panels), respectively.

The CHH participation plots first allow us to identify which normal modes contain significant contributions of CH bending character. In apo-EDTA the CH_t mode is confined to the indices 0–9 (Figure 4.5a, top), the CH_w mode is mostly prominent in the 9–19 modes (Figure 4.5a, middle), and the CH_s vibration appears most strongly in modes 15–24 (Figure 4.5a, middle). The same trend holds independent of metal binding (Figure 4.5b, c).

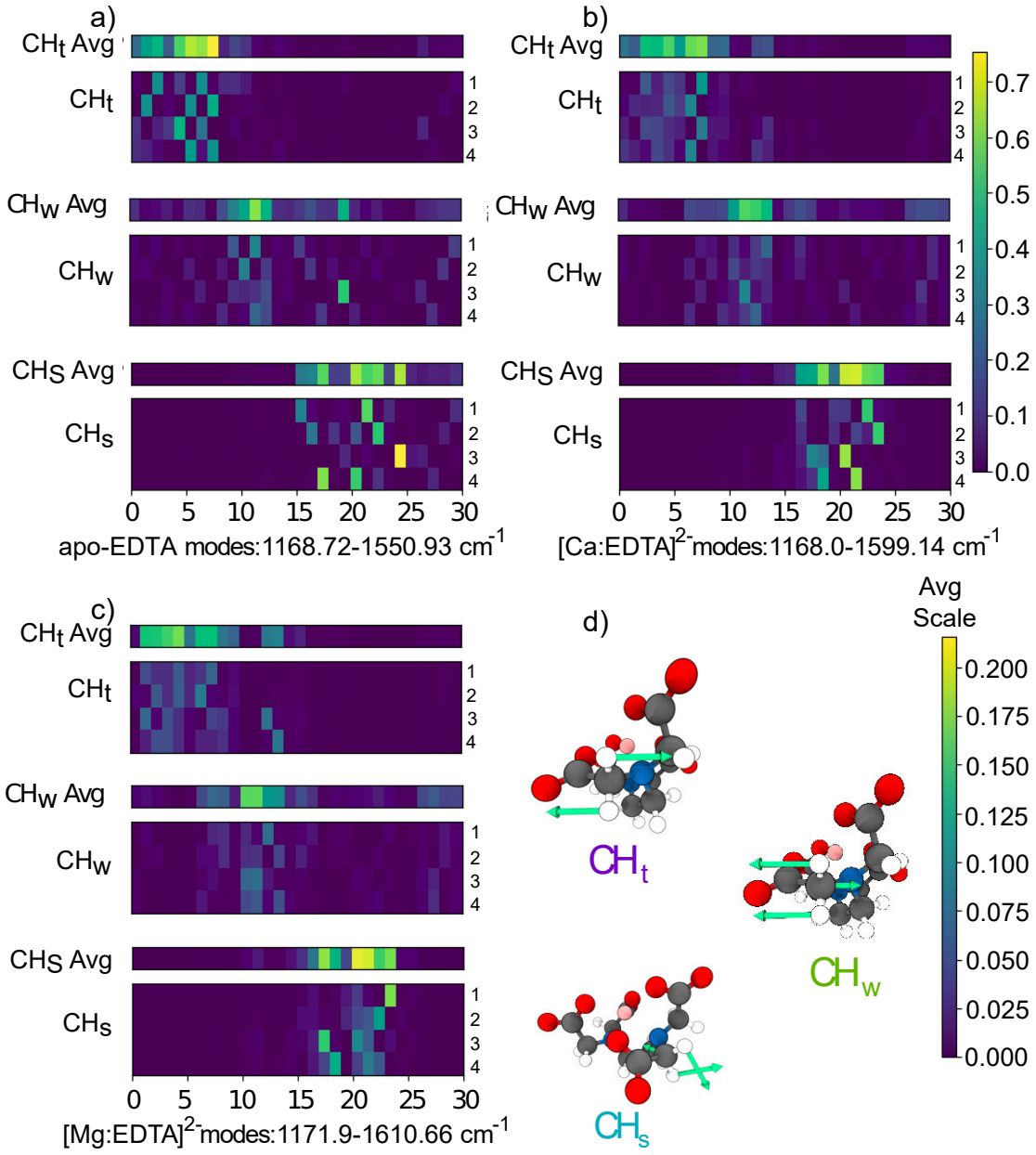

Figure 4.5: Contribution from the CH_t, CH_w, and CH_s modes to individual modes in the (a) apo-EDTA, (b) $[Ca:EDTA]^{2-}$, and (c) $[Mg:EDTA]^{2-}$ molecules differs. The four terminal acetate CHH averages are above the respective CH_t, CH_w, and CH_s participation plots. 1 and 2 are the sagittal acetates, while 3 and 4 are the equatorial acetates. Percent participation is obtained by multiplying with a factor of 100 to the scale. (d) shows the CH_t, CH_w, and CH_s motions.

For COO_a^- normal modes, the CHH participation coefficients show that they have minimal overlap with the CH_t, CH_w, and CH_s vibrations in the three molecules. Looking at indices 26–29 of apo-EDTA, we see only a minimal overlap of the CH_w modes ($< 15\%$) (Figure 4.5a, top) with the normal modes of the COO_a^- vibrations; the CH_t (Figure 4.5a, middle) and CH_s (Figure 4.5a, bottom) normal modes overlap even less ($< 1\%$). Similarly, in both $[Ca:EDTA]^{2-}$ (Figure 4.5b) and $[Mg:EDTA]^{2-}$ (Figure 4.5c), modes 26–29 have an insignificant contribution from the CHH participation coefficients. This shows that even in the complexed form, the COO_a^- do not mix with the CH_t, CH_w, and CH_s vibrations.

For COO_s^- normal modes, on the other hand, the CHH participation plots show significant mixing of the COO_s^- modes and the CH motions, especially CH_s modes. Modes 14–18 of apo-EDTA elucidate the prominent CHH participation projection in these normal modes. This region in apo-EDTA, which describes the majority of COO_s^- motion, has significant projections ($\sim$33–63%) from the ideal CH_s modes (Figure 4.5a, bottom) and minor contribution ($\sim$ 16%) from the ideal CH_w modes (Figure 4.5a, middle). From these observations, we infer that the COO_s^- mixes both with the CH_s and the CH_w vibrations. We also observe that the CH_s modes project only onto one of the CHH near the carboxylates, reflecting the localized character of these modes. Similarly, in $[Ca:EDTA]^{2-}$ and $[Mg:EDTA]^{2-}$, the COO_s^- modes mix more with the CH_s than with the CH_w modes. Therefore, we conclude that after complexation the COO_s^- mixes more with the CH_s modes ($\sim$45%) than the CH_w modes ($<$10%).

The participation coefficients allow assignments of the observed features in the linear absorption spectrum (Figure 4.6). The left column shows all the normal modes as stick spectra and with $8\,cm^{-1}$ Gaussian broadening. Modes 0–30 are assigned based on the participation coefficients. The right column presents the COO_s region (1400–1550 cm^{-1}) overlaid with the experimental absorption spectra. The wavenumber axes are shifted to accommodate the differences.

In the COO_s region (Figure 4.6b), apo-EDTA's three main peak features are the mixed COO_s^-, CH_w, and CH_s modes. The five symmetric stretches are clustered within the main peak, A ($1407\,cm^{-1}$). The most intense COO_s^- peak (mode 16), is a carboxylate stretch with participation from scissor modes. The remaining three modes are mixed with CH_w and CH_s modes. Peak B ($1429\,cm^{-1}$) encompasses the modes 20 and 21, which are CH_s motions. Band

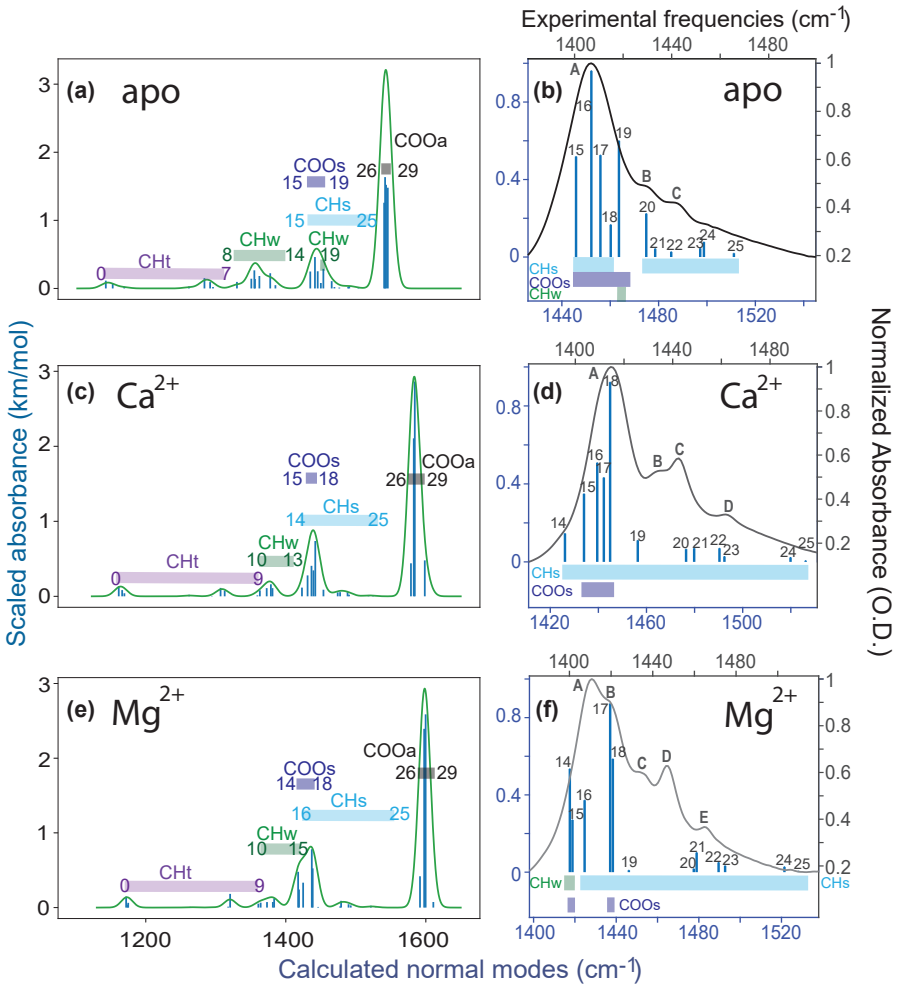

Figure 4.6: Indices 0–9 are CH_t vibration, indices ∼10–15 are CH_w modes, the 16–25 region constitutes the CH_s modes, and the COO_a^- modes are indices 26–29. The COO_s^- vibrations fall within the CH_s frequencies. The three rows represent the three molecules, the left column shows the normal modes and the experimental COO_s^- overlays the same normal mode region in the right column. Apo-EDTA is the top row((a), (b)), $[Ca:EDTA]^{2-}$ is middle row ((c) and (d)), and $[Mg:EDTA]^{2-}$ is the bottom row ((e), (f)). Spectra shown here are calculated with M05-2X/def2-TZVP/SMD theory, Gaussian width = 8 cm^{-1}, intensity scaling factor = 0.0015, and 1000–1650 cm^{-1}.

Table 4.1: Normal mode assignment from M05-2X participation coefficient (pc) analyses in apo-EDTA, $[Ca:EDTA]^{2-}$, and $[Mg:EDTA]^{2-}$ molecules for 1000–1650 cm^{-1} range.

pc index	Calculated frequency (cm^{-1})	Predominant vibration
	$[apo\text{-}EDTA]^{4-}$	
0–7	1169–1314	CH_t
8–14, 19	1346–1399, 1464	CH_w
15–18, 20–25	1446–1460, 1475–1511	CH_s
15–19	1446–1464	COO_s
26–29	1546–1551	COO_a
	$[Ca:EDTA]^{2-}$	
0–9	1168–1363	CH_t
10–13	1367–1385	CH_w
14–25	1426–1526	CH_s
15–18	1434–1445	COO_s
26–29	1580–1599	COO_a
	$[Mg:EDTA]^{2-}$	
0–9	1172–1361	CH_t
10–15	1364–1419	CH_w
16–25	1425–1528	CH_s
14–15, 17–18	1418–1438	COO_s
26–29	1592–1611	COO_a

C (1443 cm^{-1}) corresponds most closely to the CH_s vibration in modes 22 and 23. The broad and asymmetric line profile of the COO_s^- band in apo-EDTA results from the low intensity scissor motions in the modes 24 and 25 (Figure 4.6a, b). The relative IR intensity of these scissor modes are at most 20% of the mixed COO_s^- modes. We expect, based on the 2D-IR lineshapes, that disorder in the solvent broadens the frequency distribution for these scissor modes.

The absorption features of the $[Ca:EDTA]^{2-}$ symmetric stretch region are composed of COO_s^- and CH_s modes (Figures 4.6d). The main band, A (1416 cm^{-1}), contains all the COO_s^- normal modes (modes 15–18). We assign the absorption band B (1434 cm^{-1}) to CH_s vibrations (modes 20 and 21). We assign absorption bands C and D (1443 cm^{-1} and 1462 cm^{-1}) to pairs of CH_s stretches (modes 22 and 23 and modes 24 and 25, respectively).

Alternatively, the absorption band at 1434 cm^{-1} (peak B) could be mode 19 (Figures 4.6d). We disfavor this assignment because mode 19 is an out-of-plane rotation motion of the ethylene CHH and the terminal CHH in the sagittal plane, while modes 20 and 21 are the in-phase and out-of-phase CH_s vibrations of all the terminal CHH, respectively. In the CHH participation coefficients, modes 19, 20, and 21 have participation from the terminal sagittal CH_s (Figure 4.5b rows 1 and 2), which indicates that these terminal atoms contribute to the three normal modes. Taken together, the calculated IR intensity and the CHH participation coefficients of modes 20 and 21 are stronger than mode 19 (Figures 4.6d and 4.5b). Furthermore, the ωB97X-D calculation reveals that mode 19 lies within the bandshape A (1416 cm^{-1}) (Appendix A.1). Thus, the peak B is more likely modes 20 and 21 rather than mode 19.

In the $[Mg:EDTA]^{2-}$ symmetric stretching region, two of the peaks are groups of COO_s^- modes and the remaining bands are CH_s modes (Figures 4.6f). The peak A (1409 cm^{-1}) envelopes the normal modes 14, 15, and 16. Modes 14 and 15 are the symmetric stretch vibrations from the sagittal carboxylates and have participation from CH_w vibration, while mode 16 is primarily a CH_s motion. The absorption peak B (1428 cm^{-1}) consists of the COO_s^- and CH_s vibrations (modes 17, 18) from the equatorial carboxylates. Mode 19 is a CH_s vibration that overlaps peak B. The remaining peaks in the linear spectrum (C, D, and E) are assigned to modes 20 and 21, modes 22 and 23, and modes 24 and 25, respectively. All these modes 20–25 are CH_s vibrations (Figure 4.6f).

The observed infrared spectrum and the calculated mode mixing in $[Ca:EDTA]^{2-}$ and $[Mg:EDTA]^{2-}$ are a result of the binding geometry. In both $[Ca:EDTA]^{2-}$ and $[Mg:EDTA]^{2-}$, the sagittal carboxylate pair (Ca^{2+} modes 15 and 16; Mg^{2+} modes 14 and 15) is further redshifted than the equatorial pair of carboxylates (Ca^{2+} and Mg^{2+}modes 17 and 18). The difference in the frequency of these vibrations depends on the placement of the metal in the EDTA binding pocket. In $[Mg:EDTA]^{2-}$, the sagittal and equatorial COO_s modes have a greater difference in frequency causing two peaks at $1409\,cm^{-1}$ and $1428\,cm^{-1}$ (A and B) to be observed (Figure 4.6d); in $[Ca:EDTA]^{2-}$, the sagittal and equatorial COO_s modes are more similar and only one broad absorption band A is observed (Figure 4.6f).

To determine the effect of geometrical distortion on the infrared spectra, we performed another set of DFT calculations where we first optimized geometry with the Mg^{2+} or Ca^{2+} ion and then removed the metal keeping the deformed geometry intact. In the deformed geometries, the calculated normal modes show the same trend – splitting in the Mg^{2+}-bound geometry and no splitting in the Ca^{2+}-bound geometry – even without the presence of the metal. This result strongly suggests that the geometrical deformation is responsible for the greater splitting between sagittal and equatorial COO_s^- modes in $[Mg:EDTA]^{2-}$ than in $[Ca:EDTA]^{2-}$.

Our analyses resolve the assignments that were unclear in the literature. Lanigan *et al.*[148] suggested that the 1315–$1340\,cm^{-1}$ peaks in the experimental spectra are COO_s^- modes. We suggest that the experimental spectral region 1315–$1340\,cm^{-1}$ corresponds to CH_w vibrations and the 1000–$1290\,cm^{-1}$ region is a CH_t motion (Figure 4.6 left). Our analysis clarifies that the dominant mode mixing into metal-bound COO_s^- is the CH_s rather than CH_w.[149] The CH_w mixing with the COO_s^- diminishes from the apo-EDTA to an insignificant level in the metal-bound state. According to our assignments in apo-EDTA, the asymmetrical broad tail in the 1450–$1530\,cm^{-1}$ region of the infrared spectrum arises from a distribution of the CH_s modes, broadened by disorder in the solvent.

We have completed assigning the peaks in the linear spectrum for the three molecules. We found that the CH_s and the COO_s^- modes mix in $[Ca:EDTA]^{2-}$, $[Mg:EDTA]^{2-}$, and apo-EDTA; the CH_w modes mix to a lesser extent. This section also compared the delocalization of COO_a^- versus the localized COO_s^- modes and determined the cause of the asymmetric

lineshape in apo-EDTA, $[Ca:EDTA]^{2-}$, and $[Mg:EDTA]^{2-}$ spectra. We also discussed how the C_2 symmetry governs the sagittal and equatorial peak pair positions in the linear spectra of the complexes and that the binding geometry deformation is responsible for the IR peak shapes in the three molecules. Finally, we elucidated some of the unresolved assignments. In the next section, we aim to rationalize the unique coupling features between the different peaks in the 2D-IR spectra of the three molecules and link them to the binding geometry.

4.4.3 Mixed Modes Couple, Creating Cross-peaks

In this section, we will focus on the observed patterns of cross-peaks in the 2D-IR spectra and interpret them with the calculated participation coefficients. Finally, we aim to link the shape of the spectra to the binding geometry of the metal ions.

The 2D-IR spectrum of apo-EDTA reflects a disordered state, whereas those of the complexed molecules show an ordered structure. Disorder in the apo-EDTA structure (C_1 symmetry) leads to the broad diagonal bands and off-diagonal features A, B, and C in the 2D-IR spectrum (Figure 4.3a), which reflect its conformational flexibility. In the complexes (Figure 4.3b,c), however, the diagonal bands and the cross-peaks sharpen, indicating the ordering of the acetate groups around the ion.

Couplings between the normal modes cause cross-peaks in the 2D-IR spectra. Our participation analysis shows that the CHH atoms attached to any particular carboxylate participate in more than one normal mode. For example, in the CHH plot of $[Ca:EDTA]^{2-}$ (Figure 4.5b bottom), normal modes 18 and 21 both contain CH_s motions of the fourth acetate group (row 4). Any excitation of mode 18 involves the CHH atoms, which then causes a frequency shift in normal mode 21 due to the local mode (CHH) anharmonicity. In the 2D-IR spectrum of $[Ca:EDTA]^{2-}$ (Figure 4.3b), we therefore expect cross-peaks between mode 18 ($\omega_1 = 1413$) and 21 ($\omega_1 = 1442$) by virtue of the CHH atoms coupling these two normal modes, which we observe (cross-peaks A, B, C). Similarly, we can rationalize the cross-peaks in the apo-EDTA and $[Mg:EDTA]^{2-}$ 2D-IR spectra by examining the CH_s participation coefficients for each of the sagittal (rows 1 and 2) and equatorial (rows 3 and 4) carboxylates. As another example, the 2D-IR spectrum of $[Mg:EDTA]^{2-}$ (Figure 4.3c) exhibits the expected cross-peak

A from the coupling between the normal modes 17 ($\omega_1 = 1408$) and 20 ($\omega_1 = 1435$), which have participation from the same equatorial CHH atoms (Figure 4.5c bottom, row 3).

No cross-peaks appear between the strong COO^-_s diagonal bands of the sagittal and equatorial carboxylates of $[Mg:EDTA]^{2-}$ (Table 4.2). These vibrations involve the motion of different atoms, so mechanical coupling is negligible. Through-space coupling of the COO^-_s modes is likely to be small as well. EDTA's COO^-_a modes have negligible through-space coupling,[149] and the transition dipole moment of COO^-_s is half that of the COO^-_a. In addition, the transition dipoles of the sagittal and equatorial carboxylates are nearly orthogonal (Figure 4.1). Modes 17 and 18 contain in-phase and out-of-phase combinations of the equatorial COO^-_s vibrations, but the calculated frequency difference is $\sim 1\,cm^{-1}$. This splitting is less than the linewidth, and cross-peaks are also not observed.

There are cross-peaks between the $\omega_1 = 1435\,cm^{-1}$ (modes 20 and 21) and the $\omega_1 = 1452\,cm^{-1}$ (modes 22 and 23) diagonal bands (Table 4.2). Modes 22 and 23 are in- and out-of-phase motions of the three CHH's attached to each nitrogen, respectively. A terminal sagittal CHH participates in the modes 20, 21, and 22, while the other terminal sagittal CH_s participates in the modes 20 and 23 (Figure 4.5c). So, both modes 20 and 21 are coupled to modes 22 and 23 by their terminal CHH and should exhibit cross-peaks between the two bands in the 2D-IR spectrum. These anticipated cross-peaks in the $[Mg:EDTA]^{2-}$ 2D-IR spectrum are observed at C, D, and E (Figure 4.3c).

Our 2D-IR cross-peak and participation coefficient analyses suggest that the coupling between the mixed COO^-_s and CH_s modes is mechanical in origin. We projected the normal modes from a covalently bonded CHH group onto the normal modes of EDTA and its complexes (Section 4.3.4.3). In the harmonic limit, normal modes are uncoupled. Due to local mode anharmonicity, however, motion of shared atoms can cause couplings between the normal modes. Thus the cross-peaks between the mixed COO^-_s band and the CH_s bands show the effect of mechanical coupling between the carboxylates and their covalently bound CHH groups. In Section 4.4.4, we will show that charge transfer, which can contribute to cross-peak formation in a 2D-IR spectrum[29], does not play a critical role in the binding geometry.

The main takeaway from this section is that our vibrational coupling analysis of the

Table 4.2: 2D-IR diagonal peak assignments for the three molecules. Normal modes listed fall within each peakshape. The peak strengths represent the oscillator strength and the cancellation from neighboring diagonal and cross-peaks in the spectrum.

$\omega_1 = \omega_3$ (cm^{-1})	Anharmonicity (cm^{-1})	Major mode	Mixed modes	Normal modes	Peak strength
	[apo-EDTA]$^{4-}$				
1407	19	COO_s^-	CH_s, CH_w	15–19	strong
1429	N/A	CH_s	–	20–21	weak
1443	N/A	CH_s	–	22–23	faint
	[Ca:EDTA]$^{2-}$				
1413	16	COO_s^-	CH_s	14–19	strong
1435	N/A	CH_s	–	20–21	faint
1442	N/A	CH_s	–	22–23	strong
	[Mg:EDTA]$^{2-}$				
1408	19	COO_s^-	CH_s, CH_w	14–16	strong
1421	23	COO_s^-	CH_s	17–19	medium
1435	N/A	CH_s	–	20–21	weak
1452	N/A	CH_s	–	22–23	strong

Table 4.3: Binding pocket ChElPG partial charges from M05-2X.

Atom	Complex	
	$[Mg:EDTA]^{2-}$	$[Ca:EDTA]^{2-}$
Metal	1.529	1.528
N1	0.418	0.426
N2	0.399	0.432
C	0.614	0.517

2D-IR spectra suggests that observed cross-peaks are due to mechanical coupling between normal modes due to the shared atoms. The terminal CHH strongly couples the carboxylates and CH_s vibrations. Additionally, we have linked the metal binding geometry to the pattern of peaks and cross-peaks in the 2D-IR spectra. The absence of through-space coupling in the $[Mg:EDTA]^{2-}$ 2D-IR spectra is due to the small transition dipole moment and orthogonal orientation of the sagittal and equatorial carboxylates.

4.4.4 Metal–EDTA Binding is Dominated by Electrostatics but Differentiated by Ion Size

Ion size is a significant factor in the binding geometry. The CH bending modes are indirectly sensitive to this size effect, which differentiates the 2D-IR spectra of the three molecules. Our analysis, however, has not yet considered which components of the binding interaction are important for characterizing the effect of the metal binding on the infrared spectral shapes. To understand what types of interactions influence the ion–binding geometry of EDTA, we calculate the different contributions to the ion–EDTA binding energy through EDA.

The solvation energy component for SAPT is zero as those calculations were performed in the gas phase. The SAPT0 terms have been relabeled according to $\Delta E_{\text{elec}} = E^{(10)}_{elst}$, $\Delta E_{\text{Pauli}} = E^{(10)}_{exch}$, $\Delta E_{\text{disp}} = E^{(20)}_{disp} + E^{(20)}_{exch-disp}$, and $\Delta E_{\text{pol}} = E^{(20)}_{ind,resp} + E^{(20)}_{exch-ind,resp} + \delta^{(2)}_{HF}$. $\Delta E^{\text{SAPT0}}_{\text{int}}$ does not include the charge

Table 4.4: Comparison of ALMO-EDA and SAPT components. All values have units of kcal mol^{-1}.

contribution	ALMO-EDA		SAPT0 (monomer basis)	
(kcal mol^{-1})	Mg^{2+}	Ca^{2+}	Mg^{2+}	Ca^{2+}
ΔE_{geom}	22.21	9.05	136.33	96.51
ΔE_{solv}	876.94	772.90	0.00	0.00
ΔE_{elec}	−1108.80	−1027.38	−914.31	−862.97
ΔE_{Pauli}	98.79	134.57	68.43	102.75
ΔE_{disp}	5.67	−4.36	−3.14	−12.13
ΔE_{pol}	−58.56	−29.32	−189.00	−143.14
ΔE_{CT}	−11.23	−27.56	−5.22	−3.47
ΔE_{int}	−174.98	−172.09	−901.67	−818.99

In order to quantify the relative importance of electrostatic and non-electrostatic interactions in metal–EDTA binding, we performed both ALMO-EDA and SAPT calculations (Table 4.4). The qualitative agreement between the two methodologies is surprisingly strong, considering the different reference wavefunctions. The geometric distortion contribution (ΔE_{geom}) for Mg^{2+} is larger than for Ca^{2+} because the EDTA deforms slightly more for the magnesium ion, which sits deeper inside the cavity formed by the EDTA cage ($d_{\text{N-Mg}} = 222\,\text{pm}$), while calcium is only partially inside this cavity ($d_{\text{N-Ca}} = 251\,\text{pm}$).

The solvated electrostatic interaction between fragments is broken down into the gas-phase interaction (ΔE_{elec}) and the solute-solvent correction that screens this interaction (ΔE_{solv}). Despite the large solvent screening effect, the total electrostatic interaction is still larger than any other EDA component and is larger than the total interaction energy. The takeaway

transfer term due to its calculation as the difference between the dimer and monomer basis induction energies, rather than as the difference between the supersystem and constrained fragment energies (as in ALMO-EDA). However, $\Delta E_{\text{int}}^{\text{SAPT0}}$ *has* been modified to include ΔE_{geom}, which is computed at the HF/def2-TZVP/gas level.

is that binding of Mg^{2+} and Ca^{2+} to EDTA is completely dominated by electrostatics, not polarization or charge transfer.

The destabilizing Pauli exchange part of the interaction (ΔE_{Pauli}) is larger for Ca^{2+} than Mg^{2+}, which is potentially due to both the larger effective ionic radius of Ca^{2+} (100 pm) compared to Mg^{2+} (72 pm) and the absolute position of the ion with respect to EDTA.[191] The energy lowering due to polarization of fragment densities in the presence of other fragments (ΔE_{pol}) is about twice as great for $[\text{Mg:EDTA}]^{2-}$ than $[\text{Ca:EDTA}]^{2-}$. This may seem counterintuitive at first, since both the neutral and dication forms of calcium are more polarizable than the respective forms of magnesium. However, this is an energy lowering term, not an observable property like the polarizability, so they are not comparable. Both the more negative polarization and less negative total frozen density interaction ($\Delta E^{(s)}_{\text{geom}} + \Delta E^{(s)}_{\text{frz}}$) for $[\text{Mg:EDTA}]^{2-}$ indicate that the frozen density state for $[\text{Mg:EDTA}]^{2-}$ is less energetically favorable than for $[\text{Ca:EDTA}]^{2-}$, again due to the decreased ion–binding pocket distance. The energy lowering due to charge transfer (ΔE_{CT}) is one of the smallest contributors to binding, even for ALMO-EDA.

While ALMO-EDA and SAPT broadly agree, there are two categories of differences between the ALMO-EDA and SAPT results: those where the trends between Mg^{2+} and Ca^{2+} are preserved but the absolute magnitudes are quite different, and those where the trends between Mg^{2+} and Ca^{2+} are reversed or there are sign changes.

For the first category, ΔE_{geom} and ΔE_{pol} are much larger in magnitude for SAPT than for ALMO-EDA. This is because the SAPT calculations use ωB97X-D/def2-TZVP/SMD geometries and not the method native ones. SAPT0 does not account for intramonomer correlation, as it starts from a HF/def2-TZVP/gas reference wavefunction, and Hartree–Fock is expected to underestimate bond lengths and have too-small atomic densities due to its lack of (in this case) dynamic correlation[192]. If the geometries used for SAPT calculations were from HF/def2-TZVP/gas and not ωB97X-D/def2-TZVP/SMD, both ΔE_{geom} and ΔE_{pol} for SAPT should be closer to the ALMO-EDA values.

For the second category, there are two notable qualitative differences between the ALMO-EDA and SAPT results: charge transfer increases going from Mg^{2+} to Ca^{2+} with ALMO-EDA but decreases with SAPT, and for dispersion, both ALMO-EDA and SAPT display the same

trend of dispersion becoming more attractive going from Mg^{2+} to Ca^{2+}, but dispersion is actually repulsive for $[Mg:EDTA]^{2-}$ with ALMO-EDA. This is most likely due to the definition of dispersion used in ALMO-EDA, which for ωB97X-D is

$$\begin{aligned}\Delta E_{\text{disp}} = &\left(E_{\omega\text{B97X-D}}[\mathbf{P}_{\text{init}}] - \sum_{A}^{\text{frags}} E_{\omega\text{B97X-D}}[\tilde{\mathbf{P}}_A] \right) \\ &- \left(E_{\text{HF}}[\mathbf{P}_{\text{init}}] - \sum_{A}^{\text{frags}} E_{\text{HF}}[\tilde{\mathbf{P}}_A] \right)\end{aligned} \tag{72}$$

(see equation 2.8 in ref. 186), which includes some amount of correlation from the density functional and not just the empirical dispersion correction, so we expect the SAPT definition (the same as the opposite-spin MP2 correlation energy[189]) to be more formally justifiable.

Still, due to structural differences in the binding of Mg^{2+} and Ca^{2+} to EDTA, questions remain about how much of each energy interaction term is due to the electron density and size of the ion versus the different EDTA geometries. The origins of the non-electrostatic interaction terms are also unclear. To analyze this, we performed "swapped metal" ALMO-EDA and SAPT calculations (Table 4.5), where the Mg^{2+} in the $[Mg:EDTA]^{2-}$ geometry is replaced with Ca^{2+}, and the Ca^{2+} in the $[Ca:EDTA]^{2-}$ geometry is replaced with Mg^{2+}. We find that although electrostatics dominate the binding interaction and the identity of the bound ion determines the EDTA geometry, there is a nuanced dependence of the remaining decomposition terms on the ion identity and EDTA geometry (Table 4.6).

The partial charges (Table 4.3) show that ion identity is primarily a size effect and not a nuclear charge effect, since the ChElPG charges for both Mg^{2+} and Ca^{2+} in their optimal binding geometries are effectively identical. We rationalize the dependence of charge transfer being only on ion identity on Mg^{2+} having fewer orbitals that can accept electron density than Ca^{2+}. It is also encouraging that ALMO-EDA and SAPT0 are in agreement between all four ion–EDTA geometry permutations, except for the small $\Delta E_{\text{CT}}^{\text{SAPT0}}$ of $[Ca:EDTA]^{2-}$, which is likely an artifact of the definition of charge transfer within the SAPT formalism (Appendix A.1), and ΔE_{disp} becoming more favorable for Mg^{2+}-in-$[Ca:EDTA]^{2-}$ with ALMO-EDA. Placing Ca^{2+} in the $[Mg:EDTA]^{2-}$ cavity gives the expected result of increases in ΔE_{elec}, ΔE_{Pauli}, and ΔE_{CT}, with Pauli repulsion more than offsetting the favorable increase in the electrostatic and charge transfer interactions. Despite this, Ca^{2+}-in-$[Mg:EDTA]^{2-}$

Table 4.5: "Swapped metal" ALMO-EDA and SAPT results. Those columns marked with a star (*) indicate the given metal is used in the complexed geometry belonging to the other metal; for example, "Ca^{2+}*" refers to the Mg^{2+}-bound complexed geometry, but with Mg^{2+} replaced by Ca^{2+}. Otherwise, this is the same as Table 4.4.

contribution	ALMO-EDA				SAPT0 (monomer basis)			
(kcal mol^{-1})	Mg^{2+}	Mg^{2+}*	Ca^{2+}	Ca^{2+}*	Mg^{2+}	Mg^{2+}*	Ca^{2+}	Ca^{2+}*
ΔE_{geom}	22.21	9.05	9.05	22.21	136.33	96.51	96.51	136.33
ΔE_{solv}	876.94	845.16	772.90	799.05	0.00	0.00	0.00	0.00
ΔE_{elec}	−1108.80	−1001.87	−1027.38	−1178.36	−914.31	−846.66	−862.97	−963.27
ΔE_{Pauli}	98.79	38.00	134.57	342.56	68.43	23.91	102.75	278.74
ΔE_{disp}	5.67	1.35	−4.36	−5.44	−3.14	−1.69	−12.13	−20.55
ΔE_{pol}	−58.56	−29.80	−29.32	−58.41	−189.00	−149.19	−143.14	−194.92
ΔE_{CT}	−11.23	−11.58	−27.56	−47.20	−5.22	−6.41	−3.47	−13.41
ΔE_{int}	−174.98	−149.68	−172.09	−125.60	−901.67	−877.11	−818.99	−763.67

Table 4.6: Dependence of each energy decomposition term on ion identity, EDTA geometry, or both.

term	dependence
ΔE_{elec}	mostly EDTA geometry
ΔE_{Pauli}	both ion identity and EDTA geometry
ΔE_{disp}	only ion identity
ΔE_{pol}	only EDTA geometry
ΔE_{CT}	only ion identity (ALMO-EDA; SAPT slightly less so)

is still a bound complex, indicating that some combination of the components which vary most strongly with changing ion identity (ΔE_{Pauli}, ΔE_{disp}, and ΔE_{CT}) modulates the metal binding depth and, therefore, controls the EDTA geometry.

4.5 Conclusion

We have assigned the unresolved infrared peaks in the symmetric stretch of EDTA. Using electronic structure theory, we present a method to estimate the mode-mixing in the symmetric stretch region. Participation coefficients resolve the assignments for peak structures seen in the 1000–1650 cm^{-1} region for apo-EDTA, $[\text{Ca:EDTA}]^{2-}$, and $[\text{Mg:EDTA}]^{2-}$. Further analysis with participation coefficients also helps us identify the delocalization in the different CH_t, CH_w, CH_s, and COO^- modes. Energy decomposition analysis shows that electrostatics drive ion binding to EDTA, but binding geometry differences are a consequence of ion size. Five main conclusions are:

(1) We provide a method to untangle the contributions of the CH and COO modes in carboxylate molecules.

(2) Using participation analyses, the individual IR assignments are CH_t (1000–1380 cm^{-1}),

CH_w (1380–1430 cm^{-1}), CH_s (1420–1650 cm^{-1}), and COO_s (1400–1450 cm^{-1}) modes.

(3) CH mode participation in the symmetric carboxylate stretch region of EDTA makes it a region highly sensitive to cation–induced geometry changes. The ion–induced geometry distortions are reflected in the pattern of CH_s, CH_w, and CH_t participation coefficients.

(4) The observed vibrational bands in both linear and 2D-IR spectra illustrate the sensitivity of the COO_s region to structural differences between Mg^{2+} and Ca^{2+} binding. The distinct splitting of the sagittal and equatorial COO_s bands seen in linear and 2D-IR spectra correlate directly to tighter binding of Mg^{2+} compared to Ca^{2+}.

(5) EDTA strongly binds both Mg^{2+} and Ca^{2+}, primarily due to electrostatics, even in the presence of solvent screening. Size effects determine the shape of the distorted EDTA, where the depth of the ion in the EDTA pocket is likely modulated by the smaller, non-electrostatic interactions.

Changes in the linear and 2D-IR spectra are connected to the changes in metal-binding geometry. The participation coefficient method provides a new approach to elucidate the detailed information contained in the COO_s^- spectral region. We have linked the changes in metal–binding geometry to the specific patterns in the linear and 2D-IR spectra. Participation coefficient analysis, when applied to the embedded carboxylate interactions in protein active sites, might uncover a more detailed interplay between the ion and the protein[193,194] than is possible with the traditional use of amide and COO_a^- internal probes. Thus, 2D-IR spectroscopy in conjunction with participation analysis may be used to understand the selectivity and specificity of EF-Hand proteins towards Ca^{2+} or Mg^{2+} cations[48,49,52,94].

4.6 Acknowledgement

The authors thank the National Science Foundation (CHE-1454105 and CHE-195484). This research was supported in part by the University of Pittsburgh Center for Research Computing through the resources provided. S.M. thanks Dr. Zhe Ren, Dr. Thomas Brinzer, and Mr. Kai Gronborg for their assistance and many useful discussions. E.J.B. thanks cclib[195,196] for the analysis framework.

5.0 Extracting Proton Kinetics from pH Jump Experiments in Protic Ionic Liquids

In the previous chapter we used a time–resolved vibrational (2D-IR) spectroscopy along with electronic structure theory to get structural information from EDTA–metal interactions. This chapter will focus on another ultrafast time–resolved spectroscopy (TR^MPS) technique to ascertain the kinetics proton transfer in a protic ionic liquid. Here, we will examine the mechanism of proton transport in a PIL compared with water.

5.1 Introduction

Ionic liquids (IL) are solvents that are liquid at room temperature and are composed of a cation and anion pair. ILs are broadly studied for their many tunable properties.[197] Protic ionic liquids are a subclass of ILs and are of interest as solvents due to their ever-increasing applications in the recent years.[101] One their tunable properties is their hydrogen bonding. PILs have an extended, three-dimensional hydrogen bonding network like water.[97,105] This similarity suggests that the mechanisms of proton transfer might also be similar.[198] Nonetheless, in both alkylammonium and imidazolium salts of protic ionic liquids, there is no mechanistic evidence of the actual proton transfer kinetics.[96,97] Proton transfer is critical to functionality of innumerable biochemical and chemical reactions.[199–204] Understanding the proton kinetics in PILs may allow us to modulate their hydrogen bonding properties and guide the rational design of new PILs for energy fuel cells.[205]

Proton transfer from a Brønsted acid donor to a Brønsted base acceptor can result in a PIL. Previous works suggest that nano-structural organization of alkylammonium PILs result in their mesoscopic structure.[206,207] Small[104], large[208,209], and wide[104] angle X-ray and neutron[210–213] scattering studies show that these nano-domains create short as well as long range order in the PIL structures. Typically both polar and non-polar sites from the cation participate in the ordering of the amphiphilic species in 3-dimensions.[213] Dielectric relaxation

spectroscopy[214], molecular dynamics[215] and empirical potential structure refinement[198,213] techniques suggest that proton hopping is crucial to the hydrogen-bond structure formation. Currently, the structure of the hydrogen-bonding network of PILs is well understood; however, the mechanism of proton transport in these structures is still unclear. Elucidating the mechanism of proton transfer will improve understanding of PILs' fundamental interactions as well as future applications.

The Grotthuss mechanism of proton transport in water is reported to cause a rearrangement in its hydrogen-bond network.[116,119,216,217] Investigating proton transfer processes has uncovered two mechanisms by which the proton travels, Grotthuss and vehicular.[113,218] The primary mechanism of proton transfer in water is proton hopping (the Grotthuss mechanism). Unlike the vehicular (diffusion) mechanism of transport, in the Grotthuss mechanism, the proton transfer does not involve solvent molecule migration.[219] The mechanism of proton

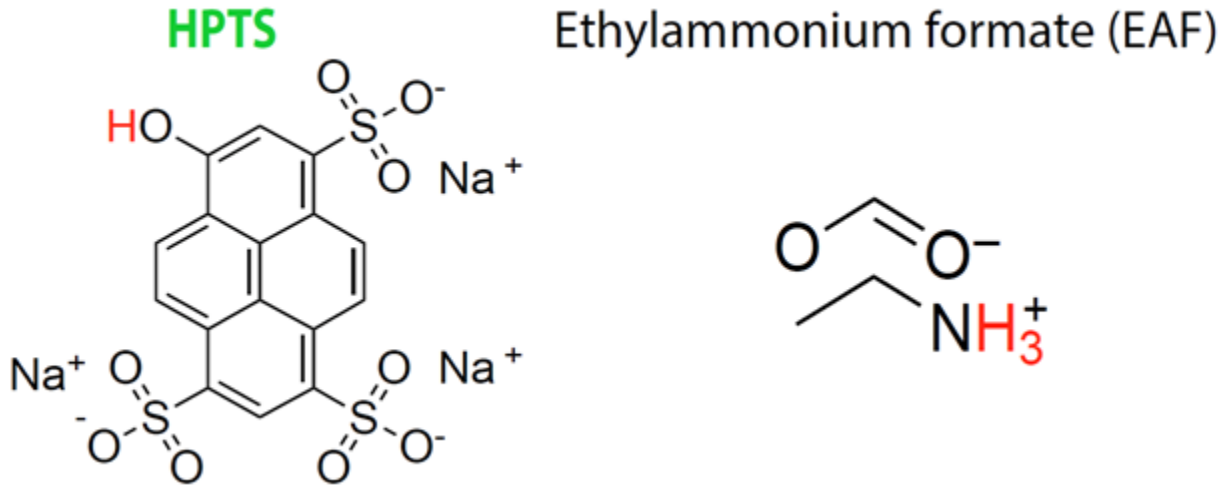

Figure 5.1: Structures of EAF and pyranine compounds.

transfer can be studied using ultrafast time–resolved spectroscopy. A pH-jump experiment can illuminate proton transfer kinetics. A ultrafast laser pulse excites a chromophore to an electronic excited state, which causes a large change in its pKa. The newly labile proton is released and transient vibrational spectroscopy can probe the kinetics. The timescales of these kinetics then provide experimental constraints on proton transfer and transport models, testing whether the proton follows a Grotthuss or diffusive mechanism.

Time-resolved multi-probe spectroscopy (TR^MPS) captures reaction dynamics from a femtosecond to millisecond timescale. TR^MPS is a pump–probe technique employing a pump-probe-probe-probe-$\cdots$ pulse sequence. The sample is excited with a pump pulse and then probed multiple times to reveal the changes in the system. Some of the major advantages of this technique includes its capability to probe the structural dynamics of a system continuously with a small sample size, procuring system information in a wide time range.[99] For proton transfer reactions, the femtosecond to millisecond measurement flexibility provides an avenue to explore both the kinetics of proton transfer at a picosecond regime, as well as the hydrogen bond network dynamics in millisecond timescales.

Photoacids are a proton source whose pK_a can be easily controlled with light. Solubilizing a photoacid in a PIL followed by a pH jump experiment supplies the labile proton we want to track. A photoacid releases a free proton into a PIL solution when excited with light of a specific wavelength. The PIL anion binds the free proton, which creates a band in the IR spectrum. In this study we use the reversible photoacid 8-Hydroxypyrene-1,3,6-Trisulphonic acid trisodium salt (HPTS), also known as pyranine. HPTS dye is extensively studied in several aprotic and protic solvents, and its pH-jump kinetics is established to be reversible and cyclic.[108,216] When 400 nm light interacts with a HPTS molecule in aqueous solution (pH 7), the $pK_a = 7.4$ of the HPTS ground state becomes $\sim$0.3 for $HPTS^*$, which ejects a proton into solution.[119,220,221] In solvents like methanol and DMSO, the vibrational bands of HPTS photoacid and photobase show negligible solvatochromic shift.[108,222] We used ethylammonium formate (EAF) as the representative PIL to study the proton transfer kinetics and infer the mechanism of proton transfer in PILs.

Both the EAF and the photoacid participate in a cyclic proton transfer mechanism (Figure 5.3). The labile proton comes from the hydroxy group in HPTS after it is excited by 400 nm UV light. In EAF, we expect five species to participate in the kinetics and hypothesize that the labile proton released from $HPTS^*$ comes back to HPTS following a cyclic reaction scheme. Thus, we predict to see the excited photobase PTS^{*-} species. We also expect the PTS^{*-} to relax back to ground state PTS^-. This should be followed by the proton being retrieved from the anion to form ground state HPTS. The overall timescale of this cyclic scheme is anticipated to be in the ps-ns regime.

Organization of this research is as follows: TR^{M}PS results on aqueous acetate solutions containing HPTS show the absorption bands of HPTS in water through the photocycle (sub-section 5.5.1). Afterwards, (sub-section 5.5.2), linear spectroscopy measurements show infrared bands of HPTS in the EAF solvent. This is followed by TR^{M}PS measurements of HTPS in EAF, which indicate the occurrence of proton transfer and the lifetimes of the different relevant species. Then we extract proton kinetics in complex EAF solvent with the help of model rate equations (sub-sections 5.5.3 and 5.5.4). Lastly, we conclude by discussing the future possibilities of improving the kinetic analysis in the EAF systems in order to better understand the complex kinetics (Section 5.6).

5.2 Materials and Synthesis

Dry HPTS was obtained from Sigma Aldrich and dissolved in EAF to make a concentration between 20-40 mM for the ultrafast experiments. Sodium acetate (NaAc), procured from Sigma Aldrich, was dissolved in water to make a 1 M solution with 40 mM HPTS dissolved in it for time–resolved ultrafast measurements. EAF was synthesized in an acid base titration. For efficient synthesis, a dry ice bath, with 5 drops of acetone for improved thermal contact, is used to bring the flask temperature to -70 °C. 70% ethylamine was added to a round bottom flask and degassed by bubbling with N_2 gas. A burette adds formic acid dropwise to the ethylamine base until an equimolar ratio is reached. It takes about an hour to get an EAF solution of 20 mL. A vacuum pump was used to dry the EAF for about six to seven hours until the EAF contained less than 5% water. EAF formation and water concentration was confirmed via Fourier transform infrared (FTIR) spectroscopy. A Schlenk line was used to dry the EAF again immediately before the ultrafast experiments.

5.3 Spectroscopic Measurements

5.3.1 FTIR & UV-Vis

FTIR samples of EAF and the aqueous solutions, were prepared in a Harrick cell setup using 25 μm spacers between two CaF_2 windows. The cells were assembled in an N_2 purged glove bag. The FTIR spectra were taken in a Nicolet 6700 (ThermoFisher Scientific) instrument at $\sim$2 cm^{-1} spectral resolution, with a constant N_2 purge during experiment. UV-Vis spectra of the same samples were taken at 1 nm resolution in a constantly N_2 purged environment. All the FTIR and UV-Vis data were collected in the Central Laser Facility, Rutherford Appleton Laboratory, STFC, UK.

5.3.2 Time–Resolved Multi-Probe Spectroscopy

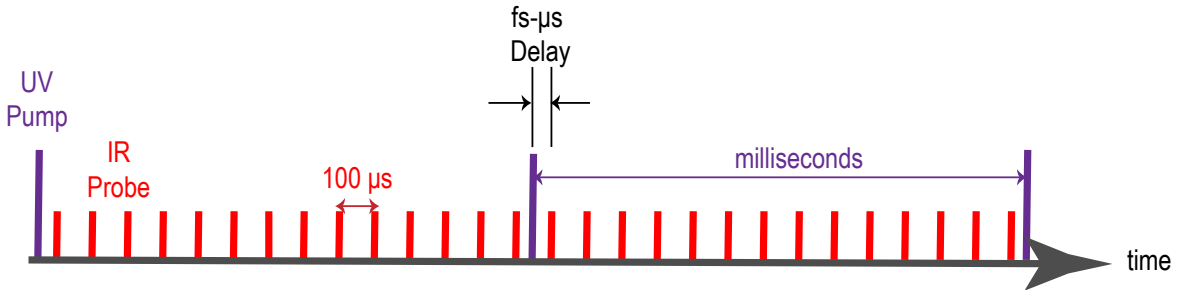

Figure 5.2: The pulse sequence in a TR^MPS experiment. The UV pump pulses (violet) have a repetition rate of 1 kHz while the probe pulses (red) have a repetition rate of 10 kHz.

The transient absorption spectroscopy experiments were performed on the TR^MPS instrument at the Central Laser Facility in the Rutherford Appleton Laboratory, UK.[99,223] Briefly, an oscillator titanium sapphire laser seeds two separate titanium sapphire optical parametric amplifiers producing the pump and probe pulses. By synchronizing the two regenerative amplifiers at two different repetition rates, the TR^MPS pulse sequence is generated. In our experiments the sample was pumped with a 1 kHz repetition rate pulse and then

probed with 10 kHz repetition rate pulses at different time delays between 1 fs to 1 ms (Figure 5.2). Our pump wavelength was 400 nm from one OPA, while the other OPA created the probe pulse centered at 1560 cm^{-1}. The spectral range for $TR^{M}PS$ was set to ∼1240–1880 cm^{-1}. The energy was 1μJ/pulse for both the pump and probe pulses. The pump OPA spot size was kept at ∼150–200 μm, while the probe OPA was limited to a tighter spot size of ∼70–100 μm. The samples were rastered to prevent photo-degradation. The probe was diffracted by a 150 lines/mm grating onto a set of two 128 element mercury cadmium telluride (MCT, IR Associates) array detectors, which relay the signal into a connected computer. All analysis is performed with MATLAB (Mathworks) software.[224]

5.4 Kinetic PreProcessor Simulation

Chemical kinetics of complicated systems can be simulated with an open source software called Kinetic PreProcessor (KPP).[225] When a set of coefficients and their respective chemical reactions are used as input, KPP can generate the ordinary differential equations for the kinetic scheme in MATLAB and other languages. The differential equations are numerically integrated over time to determine time-dependent concentrations. Taking these differential equations, one can determine the best fit parameters to the measured kinetics. The set of reactions in equations 73a–73e was used to model the photocycle of HPTS in EAF (Figure 5.3).

$$HPTS^{*} + HCOO^{-} \xrightarrow{k_1} PTS^{*-} + HCOOH. \tag{73a}$$

$$PTS^{*-} \xrightarrow{k_2} HPTS^{*} + HCOO^{-}. \tag{73b}$$

$$PTS^{*-} \xrightarrow{k_3} PTS^{-}. \tag{73c}$$

$$PTS^{-} + HCOOH \xrightarrow{k_4} HPTS + HCOO^{-}. \tag{73d}$$

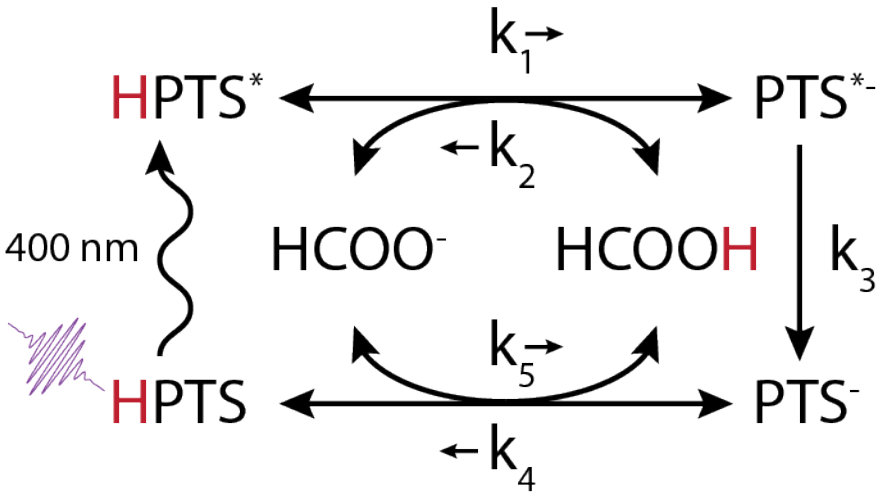

Figure 5.3: The cyclic reaction scheme used during the KPP simulation. When 400 nm UV light strikes a HPTS molecule it goes into the HPTS* excited state. The proton leaves the excited state and attaches to a formate anion in solution. Meanwhile, the excited PTS^{*-} relaxes back to ground state PTS^{-}. Finally, ground state PTS^{-} retrieves the proton from solution and equilibrates with ground state HPTS completing the cycle. Each of these forward and backward reactions are considered to have different rate constants.

$$HPTS + HCOO^- \xrightarrow{k_5} PTS^- + HCOOH. \tag{73e}$$

5.5 Results and Discussion

5.5.1 Peak Assignments in Water

To better ascertain the peaks of the different species of HPTS in a PIL solution, we first measured the FTIR spectra of 100 mM HPTS and PTS^- dissolved in water separately and then the kinetics of proton transfer of HPTS in 1 M aqueous acetate solution. These experiments were in good agreement with literature values (Figures 5.4 and 5.5).[117] The trends seen in our HPTS–acetate solution TR^MPS spectra, which shows a rise and decay of the different HPTS components in solution along with the gradual rise and decay of the acetic acid transient peak on the 10 ps timescales, agrees with previous reports.[119] We expect to observe the same bands of HPTS molecule in the EAF solution, with solvatochromic shifts within $\sim 3-6$ cm^{-1}.

In the acetate solution, the HPTS ground state bleach is at ~1404 cm^{-1}, the $HPTS^*$ transient absorption is at ~1390 cm^{-1}, the PTS^{*-} transient absorption peak is at ~1437 cm^{-1}, the PTS^- transient absorption is at ~1505 cm^{-1}, and the CH_3COOH transient absorption band is at ~1725 cm^{-1} (Figure 5.5b). Acetate symmetric and antisymmetric stretch ground state solvent peaks are at ~1400 cm^{-1} and ~1575 cm^{-1}, respectively (Figure 5.5a). The HPTS and PTS^- ground state absorptions are negligible compared to the intensity of the acetate solvent bands, hence they are not seen in the FTIR spectra. It is important to note that all the transient peaks reside on top of a solvent background in these experiments.

5.5.2 Spectroscopy

Pyranine (HPTS) has a fast nanosecond recycling timescale.[107] The nanosecond excited state lifetime of the HPTS photoacid is much longer than the expected ps kinetics of the labile proton in EAF and ensures the HPTS has returned to the ground state for the next

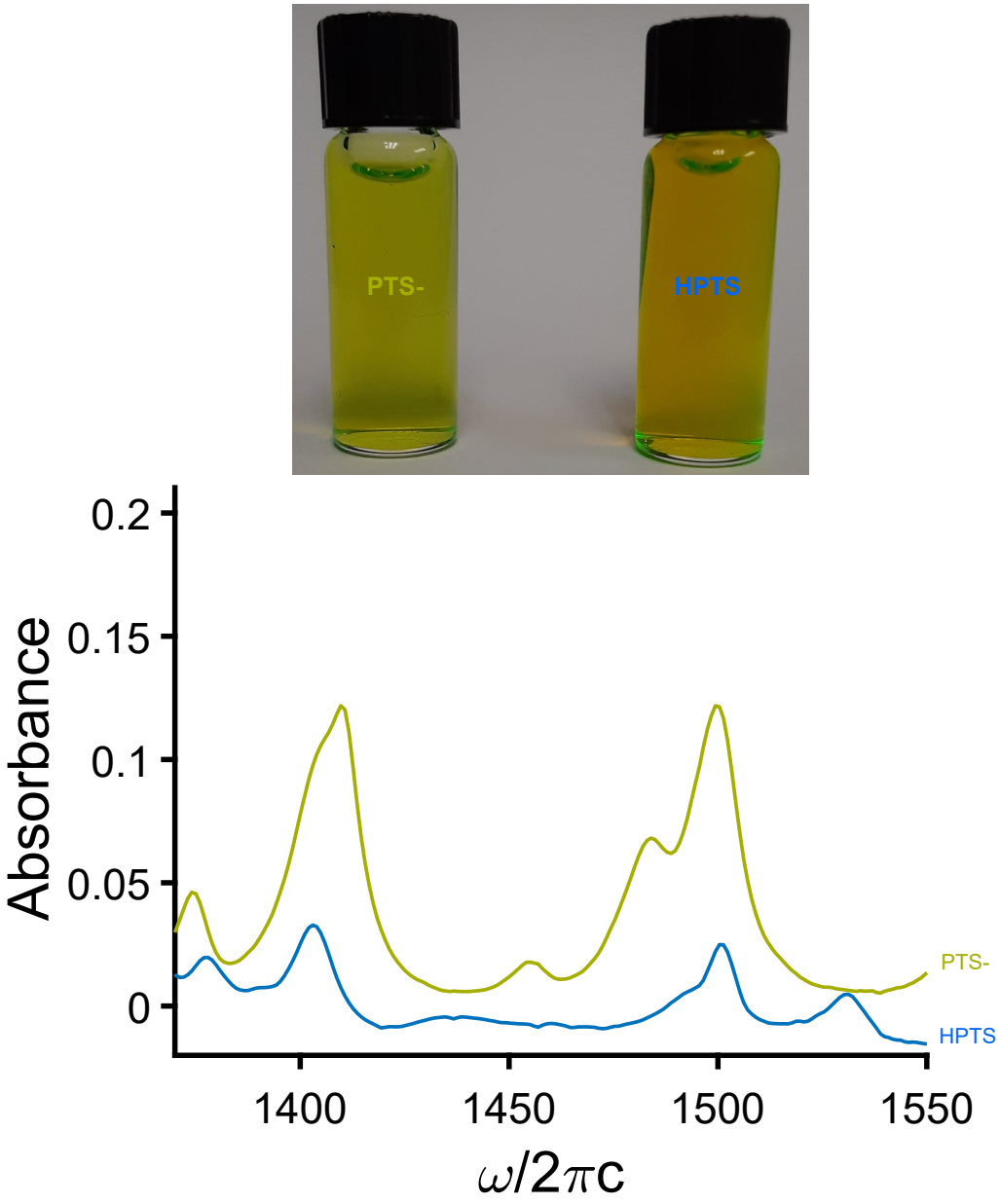

Figure 5.4: IR spectrum of 100 mM HPTS (blue) in water and 100 mM PTS^- (light green) in water. The top segment reflects the colors of the two compounds in water and the bottom segment is their respective solvent and background subtracted FTIR spectrum in the 1370–1550 cm^{-1} region.

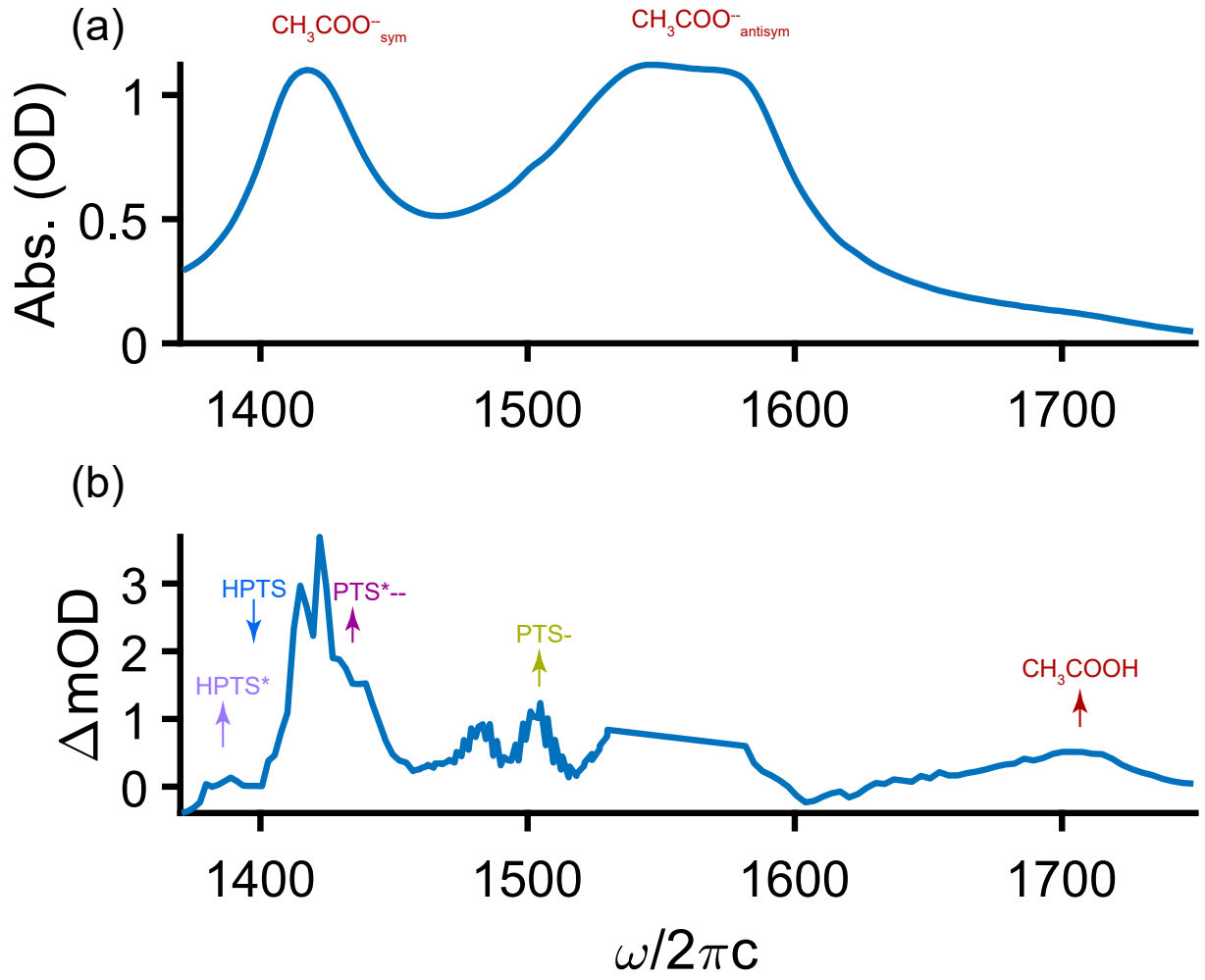

Figure 5.5: Assignments of HPTS photoacid and photobase bands for 1 M NaAc in water. The (a) top spectrum shows the FTIR while (b) the bottom spectrum is the transient absorption with a 900 fs delay after the UV pump.

laser shot. The unchanged UV-Vis spectra before and after a TR^{M}PS experiment reflects the HPTS recovery after a cycle. Also, the transparent yellow color of the solution indicated the absence of HPTS degradation (Figure 5.6).

The UV-Vis spectrum reveals the presence of a 400 nm peak designated to the excited state HPTS* species (Figure 5.6a).[221] The linear infrared spectrum shows the PTS^{-} ground state band at 1500 cm^{-1}. All the PTS^{-} bands in the FTIR spectrum are on top of the EAF solvent bands (1380–1700 cm^{-1}). The region beyond 1700 cm^{-1} has a relatively low solvent background which is utilized to monitor the growth of the proton acceptor product (formic acid) in our hypothesized kinetics scheme (Figure 5.6b). Our kinetic scheme first involves the excitation of ground state HPTS by a 400 nm light. Similar to water, HPTS* will release a proton that transfers to the solvent's formate anion generating a formic acid band in the free spectral window around ∼1700 cm^{-1}. Solvent effects, well studied for HPTS in protic and organic solvents[108], likely play a role in bringing about the 3–5 cm^{-1} solvatochromic peak shifts of the HPTS species in EAF, which is a highly ionic environment.

Time–resolved experiments reflect the proton transfer timescales in the EAF–HPTS system and a transient absorption two dimensional plot reports the ps–ns lifetime of the HPTS, HPTS*, PTS^{*-}, PTS^{-}, and HCOOH species (Figure 5.7). The red regions represent the excited and ground state absorptions of different species, whereas the blue regions are the ground state bleaches of a vibration. Here we will describe the features of the 2D TR^{M}PS plot in general and expand those features in detail in section 5.5.3. The ∼1320-1380 cm^{-1} and ∼1525-1650 cm^{-1} regions represent the intense absorption bands from the solvent. The following is observed: (1) The HPTS* excited state absorptions appear with the instrument response and decays fast. (2) The HPTS ground state bleaches also appears with the HPTS* excited state bands. (3) Other transient absorptions from PTS^{*-} and PTS^{-} species start appearing around 1 ps. (4) Formic acid starts forming at 0.001 ns. (6) The first decay occurs around 0.05 ns. (7) Next decay is at 7 ns and (8) the HCOOH appears as a broad band till 2.5 ns. From 2.5 ns, the HCOOH separates into two peaks centered at 1695 cm^{-1} and 1728 cm^{-1}, each of which decay around 70 ns. Utilizing these peak lifetimes we can model the proton transfer kinetics.

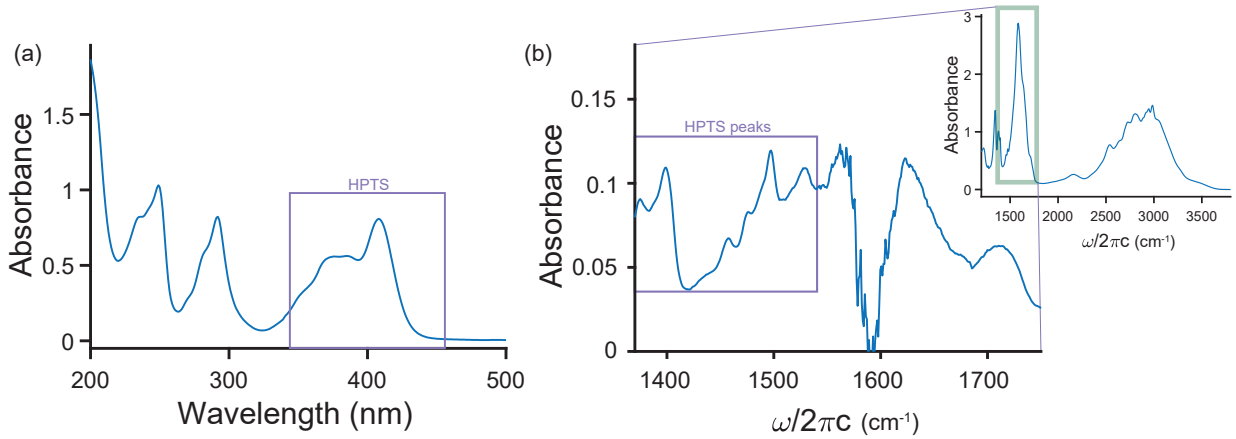

Figure 5.6: (a) A UV-Vis spectrum shows the signature 400 nm band for the HPTS* species. (b) The solvent subtracted FTIR spectrum shows the presence of bands on top of the large solvent peak in the 1370–1750 cm^{-1}. The mauve box highlights the HPTS ground state peaks in the solvent background. The inset shows the full infrared spectrum of EAF with HPTS.

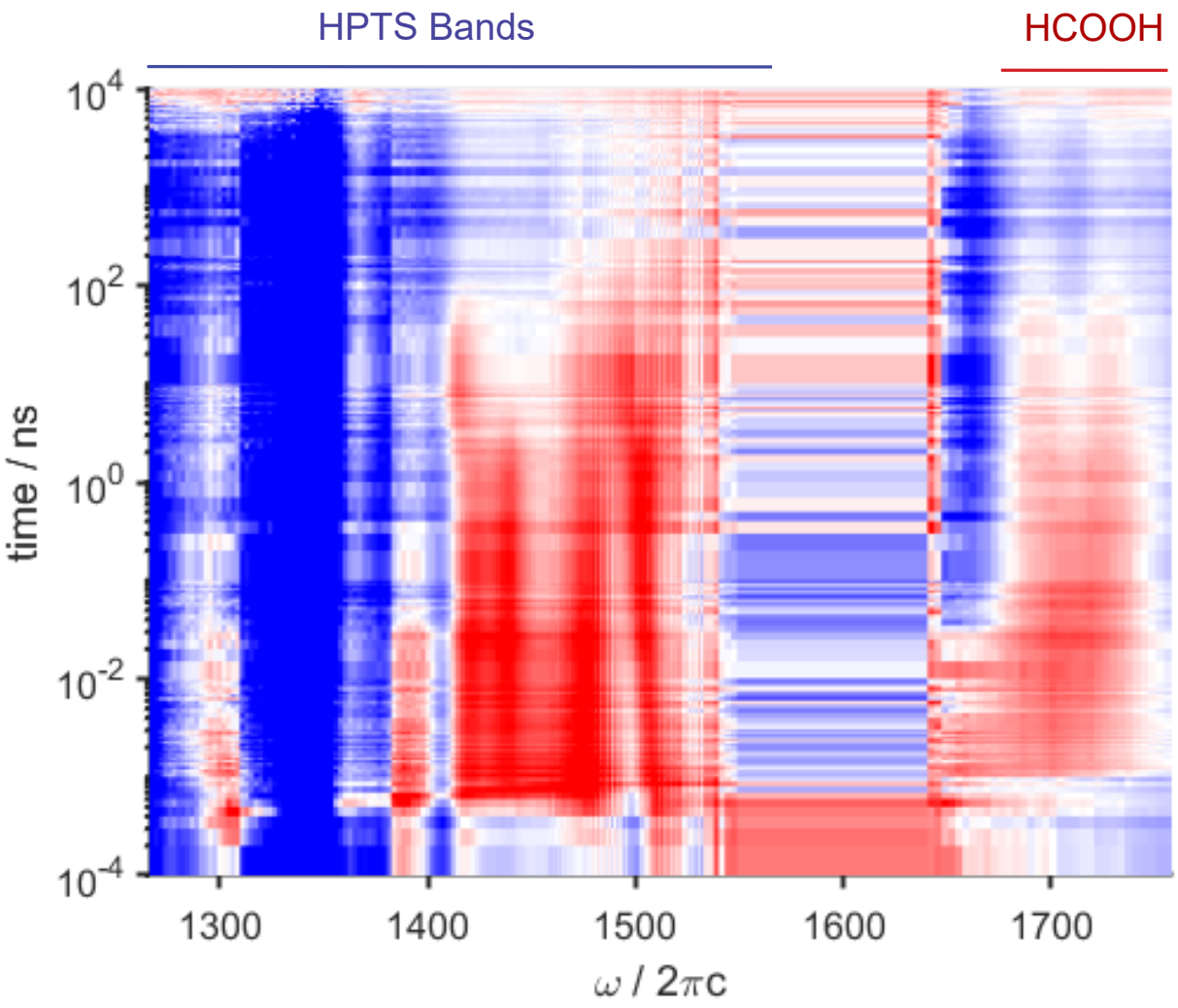

Figure 5.7: Transient absorption two dimensional plot shows the dynamics of all the species in HPTS dissolved in EAF. The blue regions are ground state bleaches and the red regions represent the excited state absorption and new absorbances. Formic acid band in the ~1720 cm^{-1} region starts growing within picoseconds of the experiment. Interspersed in this spectral window are broad solvent bands.

5.5.3 Data Processing

To extract the observed kinetics, we first processed the time–resolved data, which starts with extracting spectral shapes for the HPTS, $HPTS^*$, PTS^{*-}, PTS^-, and the HCOOH species. The HCOOH peaks beyond the 1650 cm^{-1} region was first baselined then fitted to two Gaussians (appendix A.3). For the 1350–1600 cm^{-1} range (the HPTS bands region), the characteristic spectrum for each respective species was estimated. First the solvent and the HPTS ground state peaks were located by using the difference spectra from FTIR, which agree with Mohammed *et al.*.[117] With those peak shapes the $HPTS^*$ transient bands were identified. Now we will go through the decay associated spectra (DAS) estimation process in detail below.

Five different time points show significant changes in spectral shape (Figures 5.8). (1) At time zero (Figures 5.8 t = 0 ns), the spectra shows the presence of HPTS ground state bleach and the $HPTS^*$ species. By adjusting the magnitude of the ground state bleaches a smooth DAS for HPTS was obtained (Figure 5.9 top row). Remaining positive peaks were fit to make the excited state absorption $HPTS^*$ DAS (Figure 5.9 second row). (2) Next a new shape of the spectra appears at 1.1 ps (Figures 5.8 t = 0.0011 ns). Using the broad shaped absorption at the middle of the spectrum and the same HPTS bands, the PTS^{*-} DAS is defined (Figure 5.9 third row). (3) The shape of the spectrum around 500 ps (Figures 5.8 t = 0.5 ns) helps build the PTS^- DAS (Figure 5.9 fourth row). Around this time, the broad absorption in the middle is absent indicating the presence of a new species in the reaction pool and this is taken to be the PTS^-. (4) Finally at 7 ns (Figures 5.8 t = 7 ns), new peaks rise which are set to be the *unknown* band DAS (Figure 5.9 bottom row). As of yet, the unknown bands can either be shifted solvent bands or broad PTS^- peaks, which we will discuss in depth later in this section.

Spectrum at specific time points indicate significant changes in shape, which help determine the decay associated spectra of the four HPTS species. Solvent shapes were modeled using three different regions, namely (1) the low frequency (1300–1400 cm^{-1}), (2) the high frequency (1450–1600 cm^{-1}), and (3) the dispersive shape at high frequency (1600–1650 cm^{-1}). The solvent bands have strong oscillator strength, so they were modeled by shapes that will cancel

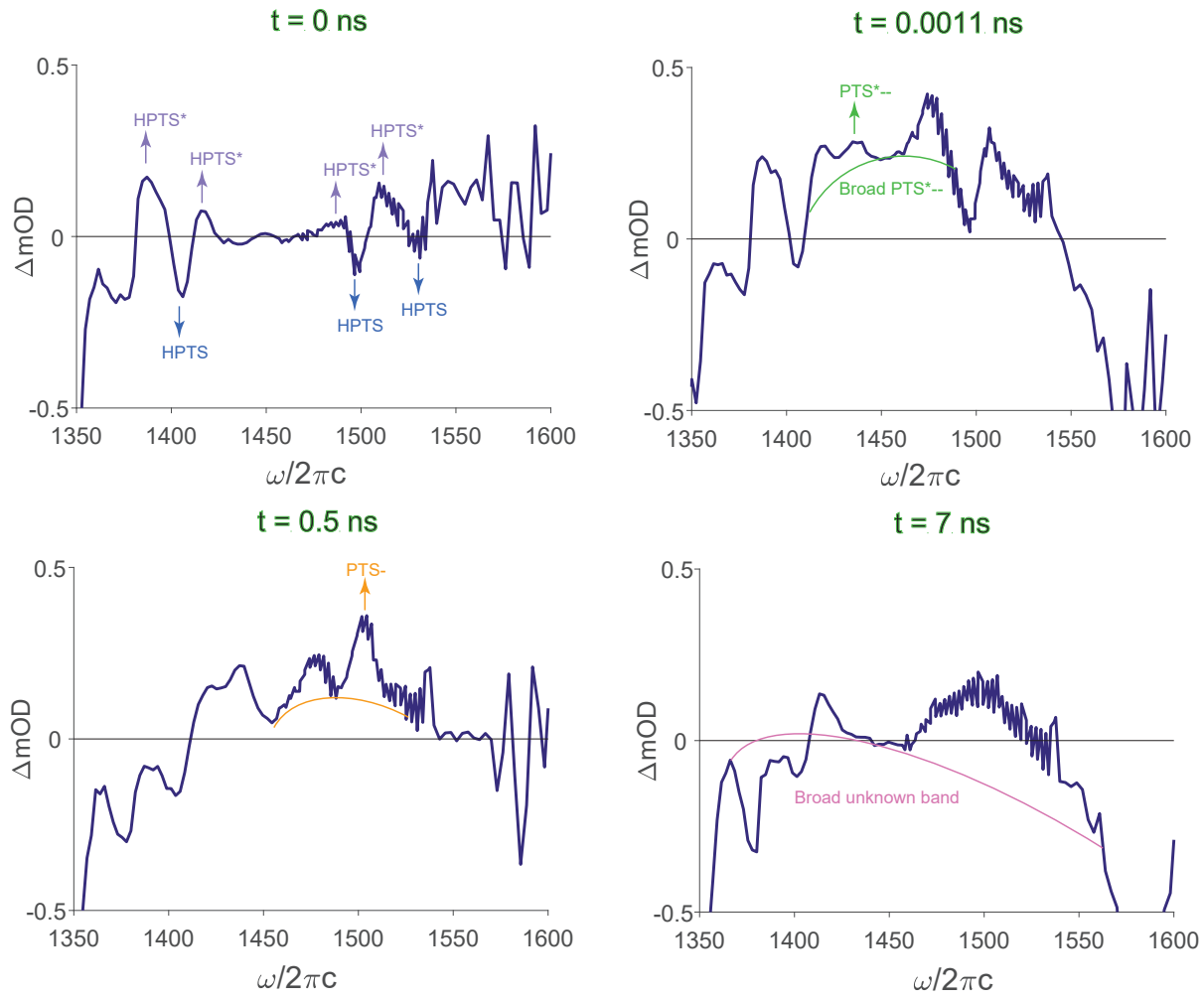

Figure 5.8: Transient peaks of the EAF–HPTS system show the rise and decay of the different intermediate HPTS products (HPTS ground state bleaches, HPTS*, PTS$^-$, and PTS^{*-} transient absorptions), the solvent bands, and the transient HCOOH peak. For each of these difference absorption spectrum, the peak shapes are distinct and useful in determining decay associated spectrum for species in the 1350–1600 cm^{-1} region.

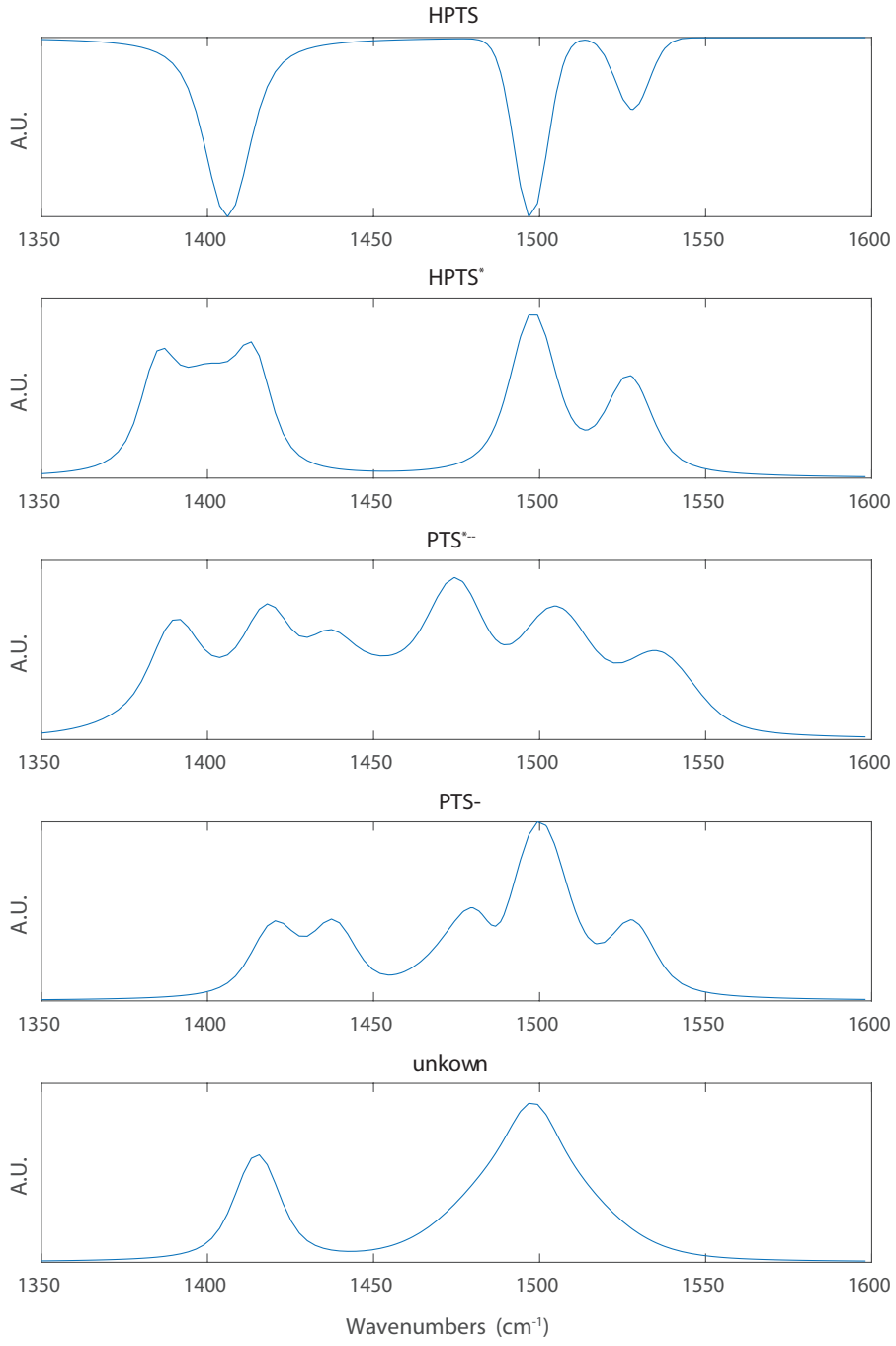

Figure 5.9: Decay associated spectral shapes in the 1300–1600 cm^{-1} region represent the HPTS, $HPTS^{*}$, PTS^{*-}, and PTS^{-} peak shapes as well as an unknown band for the solvent's thermal fluctuation. These representative shapes are fitted to the experimental data to extract the kinetic trends.

those absorptions. Following the procedure described in these two paragraphs, we outlined the five DAS for HPTS, $HPTS^*$, PTS^{*-}, PTS^-, and the unknown bands and now we will use those to draw out the kinetic trends in these species.

Applying an overall fitting function, made from each of the five DAS, we get the kinetic trend of each species. The kinetic profiles are extracted by fitting the spectral data to these functions (Figure 5.9). The amplitudes of the bands are the only variable in this fitting process. Bleaches, which are for the HPTS, can only have negative amplitudes, while transient absorptions ($HPTS^*$, PTS^{*-}, PTS^-) can only have positive amplitudes. The solvent bands can assume positive or negative amplitude.

With the DAS fitting approach, a sequential set of kinetic trends are extracted (Figure 5.10) from the 1300–1600 cm^{-1} region of the dataset. Each kinetic profile is in a logarithmic timescale starting from 100 fs and ending at 10 µs. The y-axis represents a difference absorption in mOD. First, the $HPTS^*$ species decays very fast within ∼1 ps. Even with the very low data points, the $HPTS^*$ decay shows an overall sigmoidal kinetics. This suggests that the proton transfer to the solvent happens in ultrafast picosecond timescale. Formate, which is in abundance, quickly picks up the H^+. During this time, PTS^{*-} rises quickly also indicating that proton transfer happens within a picosecond. This PTS^{*-} species decays to PTS^- within a 100 ps and this timescale of decay in the PIL is much faster than PTS^{*-} decay observed in water.[119] Future experiments need to determine if this fast decay is radiative or nonradiative in nature. Simultaneously with this PTS^{*-} decay, the PTS^- peak rises and disappears on ns timescales. When the PTS^- absorbance is zero in 10 ns, the HPTS ground state bleach is also seen to be almost recovered fully. The data points taken at very long time points, after 10 ns, have greater noise component as seen from all the species. Future experiments will aim to collect a cleaner data set at longer time points and also get more data points at earliest time points to verify the shape of the $HPTS^*$ sigmoidal decay.

The unknown band takes 10 ns to rise and decays in 100 ns (Figure 5.10 bottom spectrum). If this is a solvent DAS, then the time it stays in solution indicates the thermal relaxation of the solvent. In this scenario, the solvent cools down in ∼100 ns timescale after a temperature jump from laser pumping. An alternative hypothesis is that the unknown band is the PTS^- band, suggesting that the current DAS for PTS^- and PTS^{*-} (Figures 5.8 third and fourth

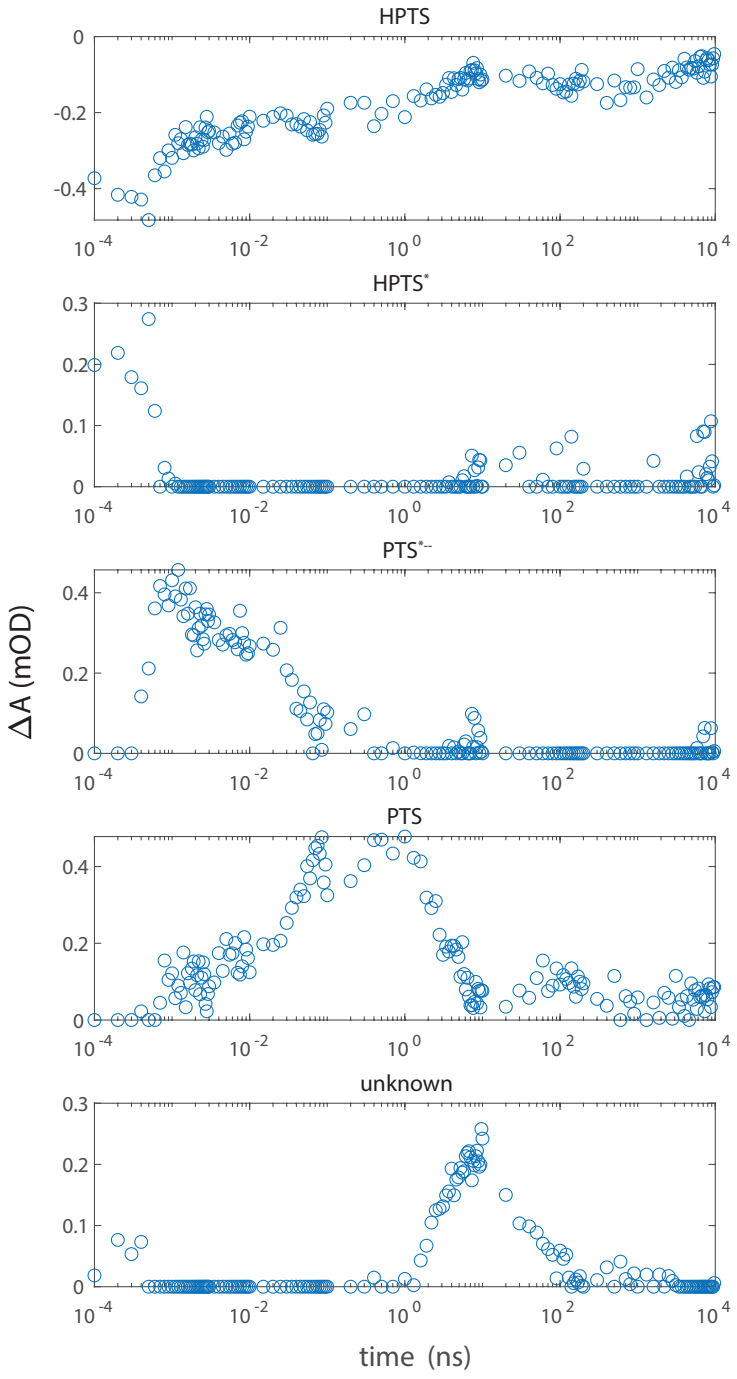

Figure 5.10: Applying DAS analysis, we have extracted the trends in the different species in the 1300–1600 cm^{-1} region of the spectrum.

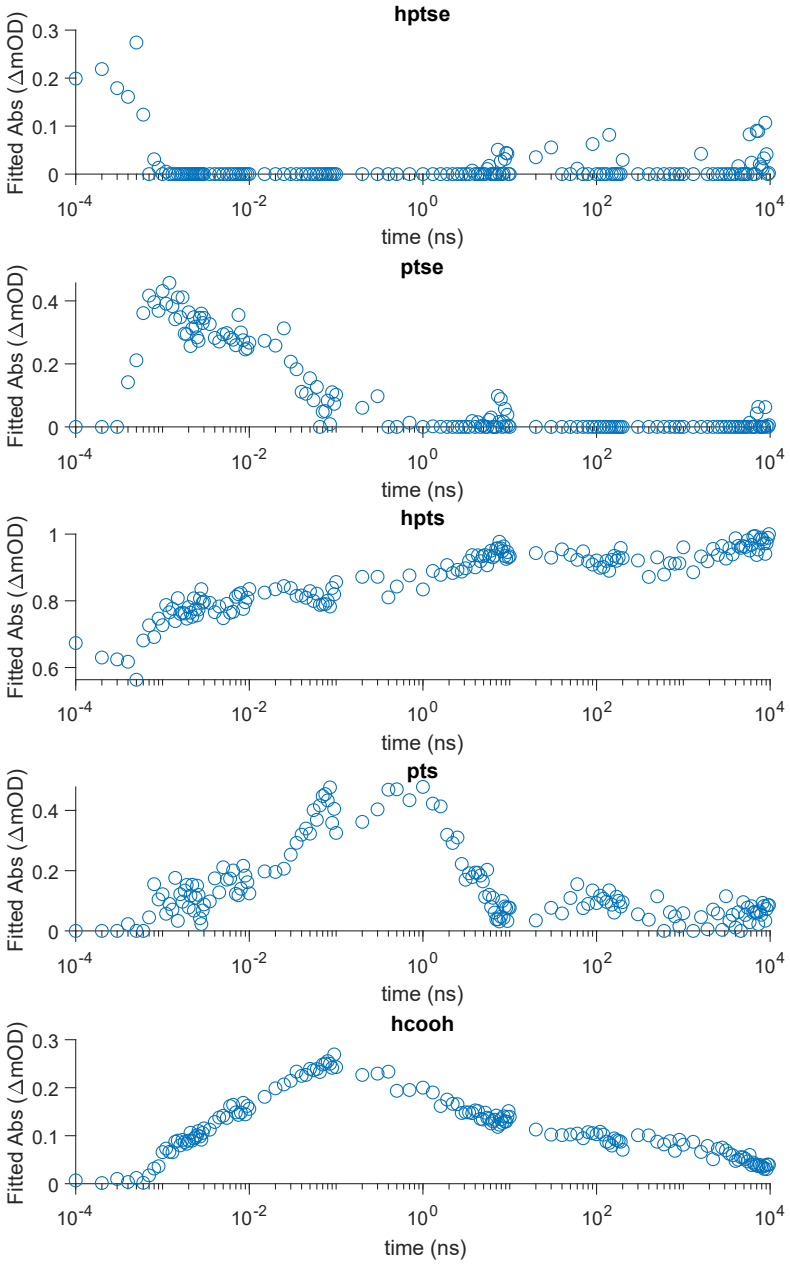

Figure 5.11: Kinetic trends (blue dots) are obtained from the DAS of the HPTS, $HPTS^*$, PTS^{*-}, PTS^- molecules. The HCOOH kinetics is from the Gaussian fitting analysis.

row) species are inconclusive. As the PTS^- molecules should to be minimum when HPTS bleaches are almost recovered (around 10 ns), we would expect an absorbance minimum in a PTS^- band; however we see an absorbance maxima in the unknown band at end of HPTS recovery. This disproves the hypothesis that the unknown characteristic shape is a PTS^- DAS and suggests that it is most likely originating from shifts in the solvent bands.

Establishing kinetic trends in the data using the DAS approach has both advantages and disadvantages. The last few paragraphs describe how this approach helps to understand the underlying kinetics in the data; however, there are caveats to using this method. Choosing specific timepoints as the starting guess for DAS can drive the model to preclude the actual data trends and force the kinetics to converge as defined by the DAS. When lineshapes are assumed at the certain timepoints, we are at the risk of imposing a kinetics on the data by virtue of starting with those artificial lineshapes. To circumvent this problem, we can use singular value decomposition (SVD) or principle component analysis (PCA) to extract the significant components affecting the kinetics and spectral shapes. In the future, a careful approach needs to be taken to eliminate the solvent background and apply SVD or PCA to draw out the accurate kinetics from the data.

For the HCOOH band located beyond the 1650 cm^{-1} region, processing included a baseline with Gaussian fitting (Appendix A.3). The presence of the large formate solvent peak in the HCOOH vicinity causes huge fluctuations in the desired formic acid transient peak. Also, we surmise that the formate solvent band undergoes stark shift due to the presence of dye molecule which leads to change in the solvent peak width. There is also a possibility of transient temperature jump[226] in that solvent band resulting from the pump pulse, during the experiment. Future experiments are needed to confirm these events in the formate solvent band. Baselining helps account for and roughly subtracts off these plausible events, and presents a relatively fluctuation–free processed peak for further analysis.

Through the Gaussian fitting, we extracted the trends of the 1692 cm^{-1} and 1725 cm^{-1} HCOOH bands (Figure 5.11 bottom). Applying a two Gaussian fit to the HCOOH band captures the overall peakshape (Appendix A.3). The two peaks (1692 cm^{-1} and 1725 cm^{-1}) symbolize the two different solvent environments that the HCOOH molecules experience in EAF. This behavior of the proton acceptor species (formate) is seen in the formate literature

when the protonated anion forms two kinds of complexes – a 'tight' complex having lower frequency and a 'loose' complex at the higher frequency.[227] Looking at the sum of HCOOH peak amplitudes we see that the kinetic trend is slow. This slow trend is a result of the initial broad peak feature (at picoseconds) separating to form two narrowing bands (at nanoseconds). Future kinetic model needs to account for the effect of peak narrowing from a broad shape to accurately extract the kinetics.

5.5.4 Kinetic Fitting Analysis

After processing, the trends in the data are fit to the cyclic kinetics generated using KPP (Figure 5.3). We estimated the upper and lower constraints of the proton kinetics for the five rates from trends seen in the DAS analysis. The best fit values were obtained by modulating the initial starting values. We plotted the difference absorbance against a logarithmic timescale to better visualize the kinetic progression. The blue dots represent the processed experimental data, while the fitting curves are the red lines in the spectra. To fit the kinetics to KPP generated equations, we utilized the linear relationship between concentration and absorbance. The KPP kinetics, which is given in concentration units, are converted to the initial absorbance inputs for the kinetic fitting process. As of yet, the PTS^{*-} extinction coefficient value is unknown in EAF (Appendix A.3). So we use the four available values and make the unknown extinction coefficient as a free parameter in the KPP fit. Finally, we convert the KPP model concentrations to absorbance units and perform the fitting. Each plot (top–bottom) is describing the evolution of individual bands of interest ($HPTS^*$, PTS^{*-}, HPTS, PTS^-, and HCOOH) in the kinetic scheme.

For the current reaction scheme, the modeled kinetics roughly captures the trends in the experiment (Figure 5.12 and Table 5.1). The fast decay rate of $HPTS^*$ exists in the fitting and so is the overall shape of the PTS^{*-} molecule well represented by the fit. The HPTS bleach recovery shows a sigmoidal shape in the fitted curve but does not fully align with the processed trend. In the PTS^-, the kinetics are better fitted to the rise of the species than the decaying side. The fitted curve needs be scaled by two in order to fit the PTS^- trend. As the estimated extinction coefficient (Appendix A.3) for the PTS^- is reasonable, the observed

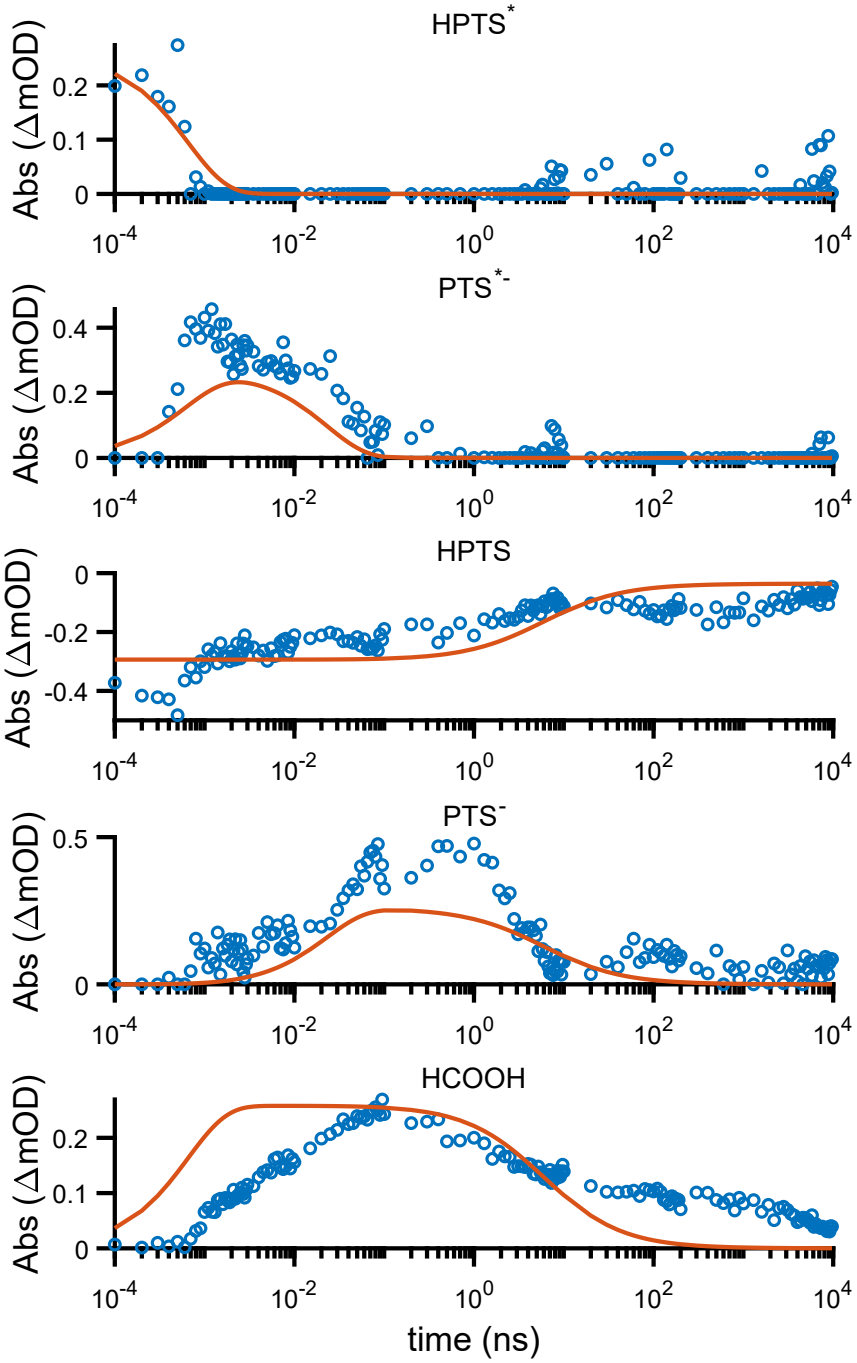

Figure 5.12: The average fits over the raw data capture the overall kinetic profiles of the five (HPTS, HPTS*, PTS*$^{-}$, PTS^{-}, and HCOOH) species. The raw kinetics is represented by blue dots while the average fitting curve is in orange.

discrepancy between fitting and the data trend is most likely from the shift in the broad feature in the PTS^- DAS (Figure 5.8, t = 0.5 ns).

The HCOOH fitting curve shows a faster rise than the experimental kinetics (Figure 5.12). These fitting kinetics come from average error minimization between the KPP and the processed experimental kinetics. The rise of the fitted curve is steeper than the data. Amplitude is missing in the data as the initial peak shape is broad causing the decreased steepness of the processed data (Appendix A.3). In the decay side (after 10 ns), the HCOOH kinetics is faster in the fit than the data. Looking at the HPTS recovery data around the same time, we see a correlated recovery trend to the HCOOH peak. This suggests that after 10 ns the HCOOH peak amplitude gets overestimated during processing due to solvent band overlap (Appendix A.3). One way to remove the discrepancies between the model and the processed data is to evaluate the standard deviation between the fits and the data. In future analysis, global fitting strategies can help minimize error between the KPP model and the experimental data.

Our current cyclic kinetic scheme does not account for the two transient peaks observed in the formic acid band. A two peak structure of the formic acid band suggests the formation of two different formic acid species (Figure 5.7). Before 10 ns, after a formate anion receives the proton, the observed transient absorption is broad. With increasing time, the broad formic acid peak bifurcates to form two bands centered at 1692 cm^{-1} and 1725 cm^{-1} and separated by ∼20-25 cm^{-1}. Therefore, in a future improvement the model needs to include the formation of these two bands. At present it is difficult to determine whether the origin of these bands are from two different HCOOH solvent shell environments or from a solvent band cooling effect.

The HCOOH band is at the edge of the solvent background (1650–1750 cm^{-1} region). So the HPTS bleach recovery is congruent with the HCOOH formation till 10 ns and then at later times (∼10 ns) the kinetics might be contributions from the growing dispersive shape of the background rather than the proton motion. The Gaussian fits conducted in these region reveal that the 1692 cm^{-1} peak undergoes frequency shifts by 3–5 cm^{-1} at longer times, but the 1725 cm^{-1} center peak frequency remains the same at all times. This might be caused due to the changing shape of the solvent background, which manifests as the 1692 cm^{-1} peak

Table 5.1: Average rates (ns^{-1}) for the five reactions.

k_1	k_2	k_3	k_4	k_5
139.1	0	43.1	0.65	0

shape and there is only one HCOOH peak (1725 cm^{-1} peak). As the background distributive shape grows in that region, we see the kinetics of the 1692 cm^{-1} peak change with time and observe kinetics of this band. The HCOOH peak fit with three Gaussians (Appendix A.3) reveal that the solvent wings at long times affect the trend seen in the data. Presently it is difficult to separate the broad solvent band shape effect from the formic acid peak shape completely. Accounting for all the solvent features accurately in future analysis is critical to the capturing the formic acid kinetics in EAF and future temperature–jump experiments on the EAF–HPTS system might provide that insight on the solvent background dynamics.

5.6 Conclusion and Future Direction

Our investigation reveals that a proton transfer reaction can be monitored using pyranine photoacid, and it essentially tracks the motion of the proton in the EAF solvent. Pyranine species peak assignments are more or less invariant in an ionic liquid system. Our results reveal that the proton transfer in EAF–HPTS occurs in picosecond timescale and slow (ns) timescale events are solvent relaxation. Kinetic analysis from our data also fits the overall shape of the actual kinetics. The current cyclic model needs to be refined to include certain HCOOH band trends. In EAF, we observe an exponential kinetics for the HPTS band where transfer of a proton to the solution is ultrafast. The broad formic acid band becoming two separate peaks in the spectra also suggests the involvement of two different formate complexes in the proton transfer mechanism.

From our kinetic analysis of the $HPTS^*$ and PTS^{*-} species in EAF, a much faster decay (radiative or nonradiative) is observed compared to water. In water, the fluorescence lifetimes of these excited state species are 4.8 ns and 5.3 ns, respectively.[119] As the lifetimes are long lived in water, they do not participate in the ultrafast reaction kinetics of the proton. In the future, fluorescence experiments of the HPTS–EAF system will describe the changes in the lifetimes of the $HPTS^*$ and PTS^{*-} molecules in a PIL environment. Probing into the average lifetime, the fluorescence quenching, and transient fluorescence experiments will give us insights into the proton transfer mechanism initiated by HPTS in a PIL. These investigations will aid in determining the exact kinetic model fit to describe the HPTS–EAF system.

Our data suggests that the proton transfer occurs fast in an EAF solution (within a picosecond). PILs are shown to have similar characteristics as water.[105,198,228] So, in order to understand this kinetics better, models analogous to water systems could be used to provide more insight. In water, the proton follows a Grotthuss mechanism of transport. Here, the solvent molecules are an integral part of the complex kinetics.[117] Considering the already established similarities between the structure of water and ethylammonium cation based PILs, it is reasonable to suggest that the kinetics in EAF proton transfer has corresponding complexity to that of water. We also observe two bands in the formate transient spectra, which suggest the existence of formates as 'tight'– and 'loose'–ly complexed in solution. Therefore, the kinetic scheme in future analysis will benefit from a model applying the pathways shown in water, where 'tight–', 'loose–', and 'solvent switch–' type complexes exist in solution.

The broad shapes of the solvent bands are complicating the accurate analysis of the proton kinetics. In addition, systematic shifts from broad solvent bands overlap with the formic acid and the HPTS bands. The DAS spectra formulated in this research were done by hand and a SVD or PCA analysis might give an unbiased method of extracting the underlying peak shapes. Preliminary analysis reveals twelve principle components that are above the noise and further analysis with SVD or PCA might show the background free kinetics. Therefore, in the future, further multivariate analysis with SVD or PCA methods may report the true kinetic trends in the EAF–HPTS system. In the future, careful temperature–jump experiments can

also examine the EAF solvent band changes in the system. *Ab initio* molecular dynamics characterization of this system will also aid in the understanding of the underlying peak shapes and the solvent behavior.

5.7 Acknowledgement

S.M. thanks collaborator Dr. Paul Donaldson at Central Laser Facility, RAL for setting up the instruments, helping with the experiment design and execution. S.M. also thanks collaborator Mr. Kai Gronborg for data analysis support and invaluable discussions.

6.0 Introducing Undergraduates to Primary Research Literature

This text is an adaptation from Mitra, S.; Wagner, E; J. Chem.; *J. Chem. Educ.* **2021**, *98, 7*, 2262–2271. The author has contributed to this work by designing, implementing, grading, assessing, and analyzing the full study, and writing of the manuscript.

6.1 Chapter Summary

All students pursuing STEM careers need to be adept at interpreting and evaluating primary research literature. The complexity of research articles requires significant practice and training in order to become skilled and efficient at this process. Thus, an early start in undergraduate education is essential and allows for development and practice through subsequent coursework and undergraduate research. The project presented here was an effort to develop fundamental literacy skills for a first semester honors general chemistry course. The curriculum included five iterative guided worksheet assignments comprising one secondary article and four related primary research articles. Each subsequent assignment was designed to be more challenging, by selecting more complicated articles and subjects, to encourage and increase literature assessment skills. Two workshops were presented after the first and third assignments to provide formal instruction on strategies to effectively read and understand research literature and promote discussion on the literature assignment articles. Assignment performance was consistently strong throughout the semester as the complexity of the assignments increased for three of the four learning objectives. However, students struggled with identifying research hypotheses despite direct instruction on this learning objective during both workshops. Pre- and post-semester survey data indicated that, while students came to the course with fairly strong confidence in their abilities to interpret and evaluate research literature, they also indicated very strong improvement in these skills by the end of the semester. These results indicate that this one-semester curriculum was an effective means to introduce students to scientific literature and develop basic literacy skills.

6.2 Introduction

Scientific training in academia should prepare undergraduates for future careers by creating curricula that develop the ability to evaluate information. A critical facet is to provide opportunities to read, evaluate, and discuss scientific literature. Survey data from our honors general chemistry (HGC) course indicated that students understand the importance of scientific literature in their careers and know that scientific developments reported in the news and secondary journal articles stem from original research. Developing the ability to effectively read and evaluate primary literature requires a grasp of article organization and writing style.[229] The complexity of primary literature warrants a scaffolded curriculum approach where fundamental skills, such as understanding the organizational design of a research article, the importance of the hypothesis, and conclusions drawn from the experimental results, are taught first. Connecting figures to the hypotheses and results, interpreting author conclusions, and a critical evaluation of the overall research design are skills that can be developed in a later segment of the curriculum. Selecting articles that correlate to the course curriculum with relevant and real-world applications is also important. The ultimate goal is to empower students to successfully apply knowledge gained from a course to interpret and critically evaluate research articles, to their own research work, and to generate new ideas.

There have been many curriculum efforts to develop undergraduate comprehension and critical evaluation of scientific literature. For example, literature database training promotes students' proficiency in sifting through journals using SciFinder, Zotero, and EndNote.[230,231] Other efforts teach researching information through a variety of journals, such as Science, Nature, and Journal of Biophysics.[232,233] These methods habituate and expose students to search methods and original research literature. Some literature curricula have students write summary paragraphs or identify key points after analyzing a primary research article,[234–236] technical report review,[237–239] or secondary journal article.[239,240]

Laboratory courses can also teach students how to interpret scientific literature and gain proficiency in generating research ideas. In methods like C.R.E.A.T.E.,[241] students follow a research group's literature or one specific topic from various research groups throughout the course and design novel experiments based on the literature. These methods increased

student cognitive skills, scientific creativity, deeper scientific understanding, and learning confidence when augmented with existing coursework. Bruehl et al. taught undergraduates the use of search engines and then to use these skills to find, review, and discuss literature on a particular topic, such as calorimetry experiments[229] Students then developed their own experiment on that topic for the laboratory component of the course. This is a great example of a scaffolded curriculum that takes students from learning to search literature to creating unique investigations. Stand-alone scientific literature courses also exist with the objective to prepare undergraduates for graduate school, independent research, choosing research endeavors, and managing both the successes and setbacks that are a natural part of the research process.[242]

The method in which educators help students improve comprehension and critical analysis of literature is important. For example, Japanese KENSHU is the scientific method of teaching students to interpret complex research and lay the groundwork for asking a set of comprehensive probing questions to test a hypothesis or draw conclusions from data and results. Drake et al. applied this pedagogy to literature assignments where students read and discussed the paper with peers one section at a time. Students wrote short summaries of the articles and presented them to the class using key figures and graphs.[243] This method was found to be effective for learning how to read and write primary research articles. Fostering critical analysis of research articles by strategically breaking down the components of a publication is also an effective educational approach.[244] Students break the article into smaller sections, analyze each part in-depth, and then join the section analyses to gain an overarching understanding and perspective of the research. Finally, students disseminate their findings in front of an audience.

Selecting research that closely aligns with course content has the added benefit of exemplifying the connection to relevant research and applications. For example, Murray used active learning POGIL methods to cover four biochemistry articles aligned with course topics.[245] Data showed that students were able to learn content as well as increase confidence in reading scientific literature. Roecker developed methodologies for an analytical chemistry course where students answered research data questions and analyzed results from articles using techniques covered in the course.[246] Another analytical chemistry course created a

scaffolded approach by dividing literature assignments into objective questions, which were completed as homework before focusing on more complicated analysis and evaluation questions in small groups during class.[247]

The multitude of publications assert that developing strong skills for reading scientific literature is of paramount importance in undergraduate chemistry education. Effective implementation of course assignments and curricula is necessary to address specific literacy objectives and improve student skills. First-year students require a broad, introductory approach to understand fundamental aspects, such as where finding hypotheses, identifying author claims, and supporting evidence. Upper level students can focus on in-depth analysis and verification of the results or critical analysis of the methods. Regardless of the class or specific goals, active learning and collaborative engagements are important in helping students understand scientific research and communication. A growth mindset is encouraged if literature is incorporated at the early stages of undergraduate education. It enables students to develop critical thinking abilities and gives them a head start on applying these fundamentals skills to undergraduate research activities and future coursework.

6.3 Curriculum Learning Objectives and Development

First semester honors general chemistry (HGC) at the University of Pittsburgh follows a traditional weekly meeting design with two lectures and a 4 h lab with the first hour as recitation. The course has a typical enrollment of 50 first-year students, nearly all of whom plan to pursue research as part of their undergraduate education. These students are focused on future careers in science, engineering, and medicine, and understand that effective reading, interpretation, and evaluation of research literature are valuable skills for their careers. Our goal in HGC is to train them for graduate level critical thinking. While some chemistry courses at the University of Pittsburgh include reading primary literature assignments, there is no focus on proficiency in these skills, especially at the first-year student level. The department does not offer a standalone literature course, which would increase the already heavy course requirement and require additional departmental resources. Instead, an

integrated approach does not increase workloads and provides the opportunity to include course activities and assignments that connect curriculum topics directly to primary research.

Development of the HGC literature curriculum was guided by the following principles. First, the curriculum includes active learning and collaborative assignments. Second, the learning process is scaffolded in an accessible manner over the two semester HGC course sequence. Third, repetition is necessary in order to build competency in the acquired skills. Finally, the literacy curriculum is integrated in the course without increasing student or instructor workload. The following learning objectives were developed for the first semester HGC course based on the four guiding principles stated above, previous efforts reported in the literature, and specific literature education needs of first-year college students. By the end of the first semester of HGC, students will be able to (1) identify primary research articles that support and correlate to concepts presented in a secondary journal article, (2) identify and explain the main goals of a research article, (3) identify research hypotheses, and (4) understand and explain fundamental aspects of the research presented, such as methodology, conclusions, future directions, and how all of these relate to the research hypotheses and goals. This first semester curriculum design is an introduction to understanding and reading literature, and it is the first step toward the scaffolded goal of a two-semester literature curriculum for the HGC courses.

Curriculum development started by selecting five broad course topics related to active research areas (Figure 6.1). Each literature assignment required one secondary research article from a journal or magazine, such as New Scientist, Discover, or C&EN news, and four primary literature references with decreasing relevance to the secondary article. Students were emailed the PDFs of the assigned articles, and the worksheets have the respective "DOIs." This made sure that the undergraduates had access to all of the resources to complete the assignments off-campus. An assignment worksheet was created (see Appendix A.4) that starts with students identifying key chemical concepts and topics in the secondary article. Next, students ranked three of the original research articles in order of correlation to the secondary article. The most closely connected primary research paper, was left out of the prioritizing list. Categorizing three original research articles with the secondary article, out of the total four papers, ensures that the most connected research article can be analyzed in-depth in

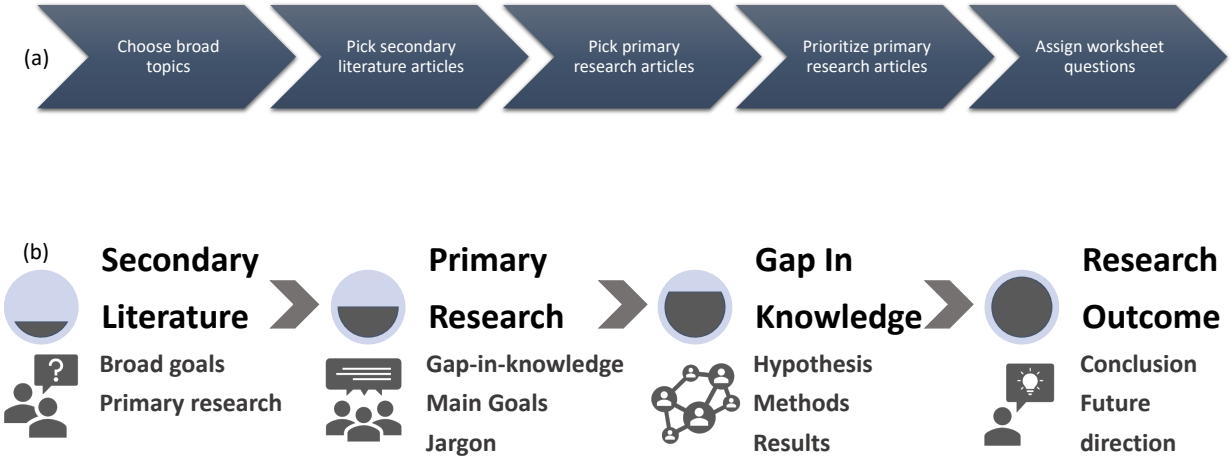

Figure 6.1: The top section (a) shows the flow of instructor preparation while the lower portion (b) shows the main topics to be addressed in the curriculum learning objectives. The secondary literature article for each assignment serves as a broad introduction to the chemistry topic and related primary research. Each assignment follows this scaffolded learning progression from broad topics presented in secondary literature to specific and detailed results and conclusions in primary research.

the later parts of the assignment. Students also learn to scan through the papers to find relevant connections to the secondary article topic, mimicking how scientists search through the literature to find relevant articles. Students pick hypotheses, write two main goals of the research, and explain their ranking to the secondary paper for two research papers. The last portion of the worksheet includes in-depth questions about the research paper that was set aside from the ranking. Students identify hypotheses and goals, deduce the research methodology in terms of "measurement,""modeling," or "making", identify conclusions and future directions, and determine whether results supported or rejected the hypotheses. As in previous questions, students quote an excerpt from the paper to substantiate their answers. In past iterations of this curriculum, we found that understanding chemistry terminology in primary research articles was an issue for students. Hence, the final version of the worksheet asked students to identify three new scientific jargon terms, define the terms, and provide a reference for the definition.

The worksheet format was purposefully designed to be generic across all five assignments to develop proficiency and skill through repetitive processes as the complexity of assigned articles and topics increased. Historically, we found that students provided long and tangential explanations when the answer length was not limited. Consequently, all answers for questions have a limited length on the worksheet template to teach students to be succinct and clear (see Appendix A.4). The primary article's complexity was based on the difficulty level of chemistry jargon and experimental methods and relatedness to the HGC curriculum. In this context, the least difficult assignment had familiar scientific terms, known or easy to interpret experimental methods, and was directly related to HGC coursework. The topics selected for the first semester HGC course were thermodynamics,[248–252] nanoparticles,[253–257] ion channels,[95,258–261] IR & MS techniques,[262–266] and carbon capture,[267–271] and were presented in this order correlating with the temporal coverage of fundamental chemistry concepts and topics in the course (Table 6.1). For example, atomic structure and spectroscopy lectures cover the basics of nanoparticles, and students performed a nanoparticles experiment in the lab component. Ion channels in biomolecules were discussed in the lecture during the broad topic coverage of drug discovery and molecular dynamics. Coverage of these topics was at a basic level and served only as an introduction to these very complicated subjects.

Table 6.1: Honors general chemistry literature assignment topics and associated articles. Topics covered in the honors general chemistry course led to the choice of literature topics and secondary journal articles. Four related primary articles were then selected for each of the five literature assignments. The last article listed for each topic is used for more in-depth questions on the assignment worksheet.

Assignment Number	Broad Topics	Specific Topics	Secondary Lit References	Primary Lit References
Lit.1	Thermodynamics	Ice Cream	248	249,250,251,252
Lit.2	Nanoparticle	Eye Floaters	253	254,255,256,257
Lit.3	Biomolecules	Ion Channels	258	95,259,260,261
Lit.4	MS and IR techniques	Fingerprinting	262	263,264,265,266
Lit.5	Environmental Chemistry	CO_2 Capture	267	268,269,270,271

Two, 30-minute active learning workshops were presented during recitation after the first and third assignments (Figure 6.2). The first workshop addressed the overall format and organization of research articles, the gap in knowledge and hypotheses, and identifying the main goals of the research. The second workshop covered advanced skills and challenging aspects for students in the first three assignments, such as determining the link between hypotheses and conclusions, identifying conclusions related to the research, and general conclusions related to the broader impact and implications of the research. During the workshop, students formed groups to discuss questions posed by the instructor related to the literature assignment submitted just before the workshop. The purpose was to use familiar literature examples so that attention could be spent on deeper comparative and critical analyses rather than attempting to read, understand, and analyze an entirely new and unknown article. The instructor facilitated group discussions to resolve differences in answers until group consensus was reached, after which the instructor brought all groups back together to state and summarize answers and facilitate overall class consensus. These two workshops kept the instructor workload at an appropriate and manageable level and while providing an appropriate amount of direct instruction to students. Regular weekly office hours for the course covered any additional help requested by students.

The literature assignments were designed to be sequentially more challenging by assigning topics and articles that were increasingly more complicated. Course lectures did not directly cover any of the literature articles but did discuss fundamental chemistry topics and concepts related to each assignment. Assignments were designed to be completed with a partner to encourage collaborative learning and reduce the grading workload for instructors. We estimated that students would require up to 4 h to complete each assignment. This includes reading requisite parts of the papers and completing the worksheet questions. To keep the course workload consistent with prior semesters, two lab experiments deemed to be the least beneficial for student learning were removed from the curriculum. Feedback from students previously enrolled in HGC indicated that each laboratory report takes roughly 10–12 h to complete after data collection. These experiments are combinations of guided and open inquiry investigations and the associated lab reports, written with a partner, requiring an abstract, short methods section, results, and a one to two page discussion created from provided writing

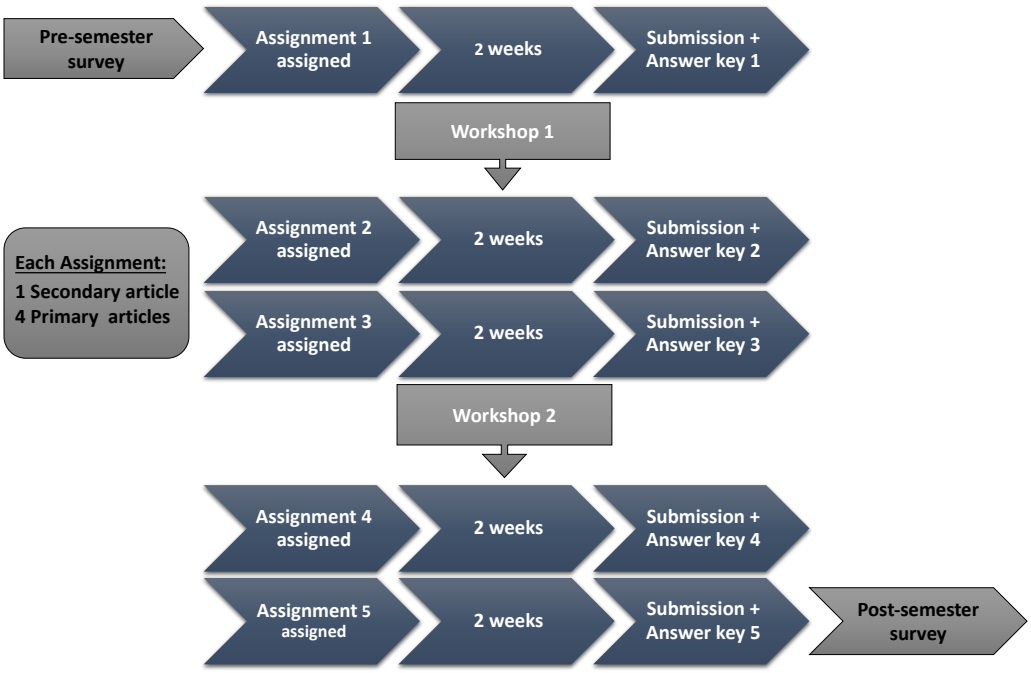

Figure 6.2: The flowchart here describes the progression of the Fall literature curriculum. A pre-semester survey, gauging the perceived literacy skills, precedes assignment 1. Five total assignments are part of the curriculum. Each assignment allows two weeks for completion, followed by submission, grading, and feedback. The workshops are scheduled after the first and third assignments to provide contextual experience on the topics covered in each workshop. The 10-week curriculum ends with a final post-semester survey about the literature curriculum.

prompts and their results. Replacing two lab experiments with five literature assignments was estimated to be an equitable time exchange that allowed students to focus efforts on developing literacy skills without overwhelming them with additional course requirements and assignments.

6.4 Methods of Curriculum Assessment

HGC is an option open to all students who wish to enroll in general chemistry at the university, however, the honors designation attracts higher-achieving students from high school. All students had at least one year of chemistry in high school, and approximately one-third had advanced placement (AP) credit for first semester general chemistry but chose to take HGC to strengthen their fundamental knowledge in the discipline. The class size was 41 first year students with broad interest in STEM majors. The course includes 3 h of lecture and one 4 h laboratory session per week where the first hour of the lab is dedicated as a lecture recitation. There were three lab sections for the course, with students evenly distributed across the sections. The curriculum was typical of a first semester general chemistry course but goes more in-depth on topics and moves at a faster pace. Infrared spectroscopy, atmospheric chemistry, and drug discovery through molecular modeling were additional curriculum topics.

Literature assignment scores, broken down into the categories that addressed the four learning objectives (stated in the introduction),were used to evaluate student skills. Specifically, questions 1, 2, 3c, and 4c addressed the ability to understand the main points of a secondary article and correlate primary research to the secondary journal article. Questions 3b, 4b, and 5b addressed the students' ability to determine the main goals of the primary articles. Question 3a, 4a, and 5a evaluated students ability to identify hypotheses, and questions 5c-5i assessed the ability to identify and substantiate research methods, conclusions, future directions, and connect conclusions to the hypotheses. A pre-semester survey collected student opinions on their knowledge and confidence with reading and understanding primary research (Table 6.2). All questions were statements about confidence and ability on specific aspects of literature and were Likert scaled from 1, strongly disagree, to 6, strongly agree. A

post-semester survey presented the same questions but phrased in terms of perceived gains in skills rather than an absolute Likert scale ranking (Table 6.3). Finally, post workshop surveys, administered at the end of the respective recitation, gleaned information specifically about the value of each workshop. These surveys include one Likert scaled question rating the helpfulness of the workshop and two open-ended questions clarifying the rating. The surveys were purposefully short to fit within the recitation period and limit survey fatigue. The same instructor was assigned to teach all three sections of the recitation and laboratory portions of the course. This instructor presented all literature workshops and graded all literature assignments.

6.5 Results

Performance on the five assignments remained strong over the semester as complexity of subject and associated articles increased. Figure 6.3 shows the assignment scores broken down into the first four learning objectives delineated in the introduction. Student perception of assignment difficulty (collected on the post-semester survey) correlates well with the intentional plan to sequentially increase complexity, except for assignment four. Each worksheet score (Figure 6.3) was kept at 25 points, encompassing all four objectives for an assignment. The class average for the five assignments were as follows; Lit 1 = 90.0% (range = 77–99%), Lit 2 = 89.6% (range = 78–97%), Lit 3 = 90.4% (range = 78–98%), Lit 4 = 88.0% (range = 78–96%), Lit 5 = 91.2% (range = 80–98%).

Table 6.2: Pre-semester student perception of literacy skills.The pre-semester opinion survey asked students to provide a response to each of the nine statements about their confidence with specific tasks related to reading and comprehending chemistry research literature. All questions used a Likert-type scale having a range of 1–6 with these categories: 1–Strongly disagree; 2–Disagree; 3–Somewhat disagree; 4–Somewhat agree; 5–Agree; and 6–Strongly agree.

Statement for Student Response: "I am confident in my ability to..."	Mean (N=41)	Median
1) read and understand secondary literature articles about chemistry.	4.7	5
2) read and understand primary research articles about chemistry.	3.9	4
3) find primary research articles related to a secondary literature article.	4.1	4
4) determine the author's hypothesis in a primary research article.	4.5	5
5) determine the meaning of new terms in the primary research article.	4.2	4
6) determine the author's goals in a primary research article.	4.2	4
7) determine whether a primary research article in chemistry is about "measuring", "making", or "modeling".	4.2	4
8) determine the conclusions drawn by the author in a primary research article.	4.4	5
9) determine the future directions suggested by the author in a primary research article.	4.0	4

Assignment one on the texture of ice cream focused on thermochemistry and how viscosity changes with molecular size and temperature. These concepts were easy for students to understand, and they did not express concerns about the assignment difficulty, although the scores for objective 3 were much lower than for the other three learning objectives. The second assignment scores declined a modest amount, most notably with identifying and understanding the hypotheses, which continued to be the lowest of the four learning objectives. Students expressed difficulty with the discipline specific terminology and scientific methods presented in the primary articles. Students expressed in the workshop and at office hours that the main goals and the hypotheses on the third assignment were very difficult to figure out, even though they performed well in the end. Interestingly, the objective 3 scores increased while the score for identifying the main goals of the research declined the most. Workshop two, which used the third assignment for discussion and question examples, allowed students to resolve the complexities they faced in assignment three. Based on both verbal and written comments, students found the writing structure, methods, and research concepts intricate, thus requiring more time to complete the assignment than the previous two assignments.

Topic complexity is intertwined with the perceived difficulty of reading these articles. Therefore, untangling the extent to which the writing style and topic each affect the perceived difficulty of the third assignment is complicated. Some insight into this question is gleaned from the fourth and fifth assignments, which were after the second workshop. Even though the difficulty of the papers and chemistry topics was anticipated by instructors to be progressively more complicated than the previous three assignments, there was a recalibration and reduction on perceived difficulty. It appears that workshop two was beneficial to students and ultimately helped them complete these assignments with less difficulty than expected. This data indicates that the efficacy of the scaffolded curriculum structure with workshops bolsters student literature reading capabilities through the semester. While students perceived an increase in assignment difficulty, overall scores remained strong for learning objectives one, two, and four. However, objective three scores indicated continued difficulty with identifying hypotheses.

The overall median Likert score on the pre-semester opinion survey was 4 out of a maximum of 6, indicating that students generally felt confident in their abilities to perform

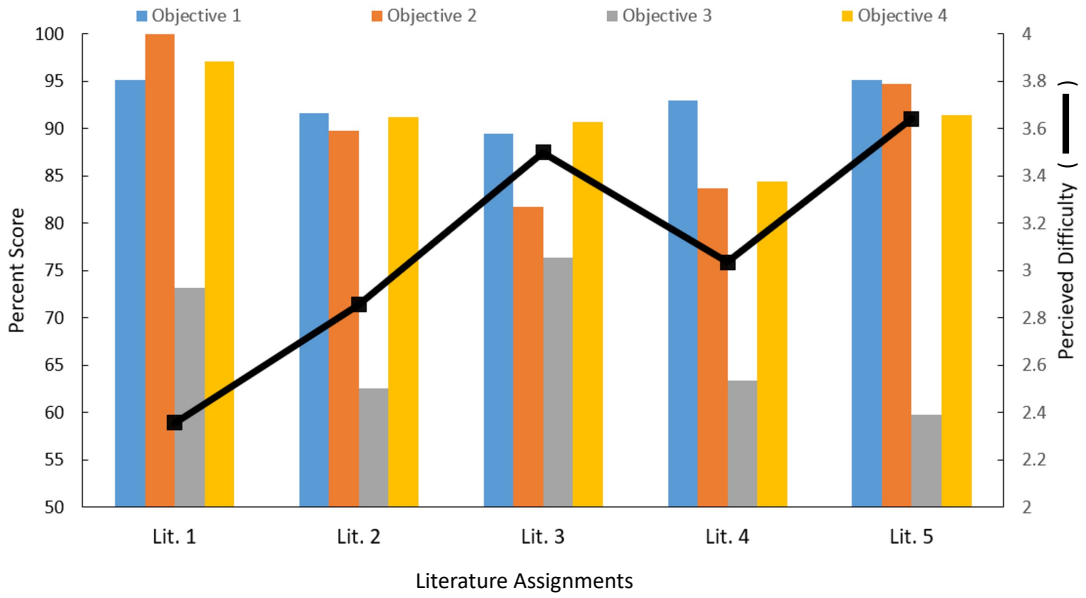

Figure 6.3: Literature assignment scores are shown in temporal order and are broken down by the four learning objectives: 1) correlating primary research articles to a secondary journal article, 2) Identify and explain the main goals of a research article, 3) Identify hypotheses, and 4) understand and explain fundamental aspects of the research presented. The class average is on the left vertical axis. Student perceived difficulty (represented by the black line) for each assignment is on the right vertical axis, with Likert scaling from 1-not difficult to 6- high difficulty. The two literature workshops were presented after assignments one and three.

the specific task addressed in each question, such as "I am confident in my ability to determine the author's hypothesis in a primary research article" (Table 6.2). The highest average was on reading secondary journal articles, while the lowest average was on reading primary articles on chemistry research. Students indicated slightly higher confidence, albeit insignificant, in the ability to perform specific reading tasks identified in questions 4–9.

Table 6.3: Post-semester perceived literacy skills improvement. The post-semester opinion survey asked students to provide a response to each of the nine statements about their perceived improvement on specific tasks related to reading and comprehending chemistry research literature. All questions used a Likert-type scale having a range of 1–5 with these categories: 1–No improvement; 2–Minimal improvement; 3–Moderate improvement; 4–Good improvement; and 5–Major improvement.

Statement for Student Response to This Question: "Compared to the beginning of the semester, how much do you feel your confidence and ability has improved in being able to perform the following tasks?"	Mean (N=41)	Median
1) read and understand secondary literature articles about chemistry.	4	4
2) read and understand primary research articles about chemistry.	4.2	4
3) find primary research articles related to a secondary literature article.	3.8	4
4) determine the author's hypothesis in a primary research article.	3.7	4
5) determine the meaning of new terms in the primary research article.	3.9	4
6) determine the author's goals in a primary research article.	4.1	4
7) determine whether a primary research article in chemistry is about "measuring", "making", or "modeling".	4	4
8) determine the conclusions drawn by the author in a primary research article.	4.2	4
9) determine the future directions suggested by the author in a primary research article.	4	4

The end of semester survey (Tables 6.3, 6.4) provides some insight into student perception of literacy skill development. This design differs from the typical pre-and post-treatment comparative assessment approach, which would have required students to accurately remember their opinions from the beginning of the semester as they complete the post-semester survey. Instead, the approach of surveying perceived improvement at the end of the semester provides a more accurate indicator of increased skills over what they indicated at the beginning of the semester. Improvement in identifying an author's hypothesis scored the lowest on the survey, which aligns with the pre-semester survey result where students indicated this to be their best skill and, thus, felt they did not need as much help to improve on this aspect. The pre-semester survey also indicated high student confidence for the ability to "determine the conclusions drawn by the author in a primary research article." However, the post-semester survey indicates this to be one of their biggest areas of improvement, suggesting that the curriculum did a better job in developing this skill than identifying hypotheses. It may also indicate that students became aware of their deficits in this skill during the semester and actively sought to improve it. Regardless of the reason, students indicated this curriculum to be a valuable and worthwhile part of the course (Table 6.3). In the post-semester survey, students were also questioned if the time spent to complete each assignment was appropriate. On a range from 1 (too little time) to 10 (too much time), the class average was 5.7, indicating that students felt that the time requirement was reasonable and that the level of difficulty was appropriate.

Table 6.4: Perceived educational value of the literature curriculum. The post-semester survey asked students to provide opinions on the overall value of the literature curriculum. All questions were scored using a scale with a range of 1 (Definitely do not agree) to 10 (Definitely agree).

Statement for Student Response	Mean (N=41)	Median
1) Compared to the beginning of the semester, I feel that I am better at finding the correlation and relatedness of primary research articles to a secondary journal article.	8.1	8
2) Compared to the beginning of the semester, I feel that I am more skilled in my ability to read and interpret primary literature read and understand primary research articles about chemistry.	8.4	8
3) The literature assignments contributed to helping me develop skills and abilities to effectively read primary literature.	7.2	7

Table 6.5: Comparison of workshop 1 and workshop 2 opinion survey results.

Questions for student response: Workshop 1

1) How helpful was today's workshop? (N = 40) mean = 7.8 median = 8

2) What specific parts of today's workshop were most helpful?

– identifying the hypothesis: *location, writing structure*

– keywords to distinguish a hypothesis from main goals: *phrases, differences*

– progression of hypothesis to main goals: *knowledge gap to story structure*

3) What would you like help with, in terms of reading primary research literature, during the second and final workshop?

– *more instructions and examples on identifying hypothesis and main goals.*

– *how to efficiently break down all the information in a scientific paper.*

– *examples demonstrating the difference between primary and secondary literature.*

Questions for student response: Workshop 2

1) How helpful was today's workshop? (N = 38) mean = 7.3 median = 7

2) What specific parts of today's workshop were most helpful?

– *linking the hypothesis with the conclusion.*

– *picking the general and specific conclusions of a paper.*

– *structure of well written sections of a scientific paper.*

3) What specific parts or topics of today's paper were not helpful?

– *determining the existence of more than one hypothesis.*

– *instructions for picking out unclearly written hypothesis or conclusions.*

– *critiquing the authors conclusion is an unnecessary exercise.*

The two workshops presented during the semester were viewed as a very helpful aspect of the curriculum (Table 6.5). The first post workshop survey included free response questions asking about the most helpful aspects and were useful topics for the second workshop. The top three most prevalent responses to each of these questions are provided in Table 6.5. This data indicates that the most beneficial aspects were understanding the organization of scientific papers, conceptualizing hypotheses, and identifying primary objectives to address the knowledge gap. Students were curious to see more examples of identifying hypotheses and main goals in primary articles. They wanted more examples of fundamental differences between secondary and primary literature. The second post workshop survey free response answers indicated that connecting the hypothesis and conclusions was most helpful. Interestingly, additional instructions to identify unclearly stated hypotheses and conclusions were not viewed as helpful.

6.6 Discussion

The overall performance on all five assignments revealed the efficacy of the scaffolded assignment approach with workshop support. The importance of the workshops was most clearly exhibited through the perceived assignment difficulty trend that decreased immediately following workshop two. Scores for learning objectives 1, 2, and 4 remained relatively consistent even though the complexity of the assignment subjects and papers increased. The continued difficulty in identifying and understanding hypotheses (learning objective 3) through all assignments is perplexing. Students come to the course with inherent confidence in this skill yet perform the lowest in this area. Understanding and identifying hypotheses was a formal part of the first workshop's agenda and was also a prevalent discussion topic in the second workshop. While students indicated high confidence in other skills on the pre-semester survey, the curriculum exposed and made them aware of the related deficits, and they worked to improve these skills. This was not the case with identifying hypotheses. The skill appears to be more complicated than anticipated and warrants additional formal instruction.

Implementing an effective and manageable scientific literacy curriculum can be complicated

and time-consuming. Hence, pre-semester planning and curriculum development are of paramount importance. In our case, two years of iterative revisions were taken to develop this curriculum. The scaffolded and generalized template of the literature worksheets allows for use with any topic. While the worksheet was designed to use four primary articles with each secondary journal article, this number can easily be revised as deemed appropriate. An important inference from this investigation is that selecting broad topics aligned with the course curriculum provides opportunities to extrapolate fundamental chemistry concepts to research and relevant applications.

Literature topics, especially for selecting secondary articles, should also align with the general interests of students. Contemporary science topics prevalent in secondary journals and news media garner the biggest student interest. In our case, the assignments with secondary articles on ice cream, nanoparticles, and carbon dioxide sequestration were most intriguing to students. Carefully selecting research articles with an appropriate and discernible level of connection to the secondary journal article is important for addressing learning objective one and promoting interest.

Finally, research articles need to be accessible at a level that supports literacy skill development and allows students to gain experience without becoming overwhelmed and frustrated. This aspect proved to be the most challenging during development. For example, in a past iteration of this curriculum, assignment scores were quite low, and survey data indicated much difficulty and frustration with understanding the research articles. In this prior version of the curriculum, seven primary research papers were included to rank relatedness to the secondary journal. The intention was for students to quickly skim the articles as scientists often do when conducting literature searches. However, student proficiency in this skill was much lower than expected, resulting in more time than intended to rank the articles.

Past iterations of the worksheet also included in-depth questions, such as interpreting figures and graphs and correlating them to the hypotheses, results, and conclusions. Students found these more in-depth analyses to be quite difficult and timeconsuming. Results show that the final version of the assignment worksheet was rated to be educational while maintaining an appropriate amount of time to complete an assignment. Student and instructor workloads should also be considered when determining the number of assignments. We found that

replacing two lab experiments with five literature assignments provided an equitable balance of work.

6.7 Conclusion

The ability to read, evaluate, and discuss scientific literature are invaluable skills for STEM majors. Providing learning opportunities early in college careers allows students to gain experience and proficiency in applying these skills through subsequent coursework and undergraduate research endeavors. The complexity of primary research articles requires a scaffolded and accessible curriculum approach where basic skills are addressed first. The project presented here was an effort to develop the following fundamental literacy skills for a first semester honors general chemistry course: (1) identify primary research articles that support and correlate to concepts presented in a secondary journal article, (2) identify and explain the main goals of a research article, (3) identify research hypotheses, and (4) understand and explain fundamental aspects of the research presented, such as methodology, conclusions, future directions, and how all of these relate to the research hypotheses and goals. The curriculum included five iterative assignments, each of which included one secondary article and four related primary research articles, and each subsequent assignment was designed to be more difficult by selecting more complicated articles and chemistry subjects. A guided worksheet was used to investigate these articles and develop the learning objective skills. Two workshops were presented after the first and third assignment as a means to provide formal instruction and discussion. Assignment performance was consistently high throughout the semester, except for identifying hypotheses. Students continued to present difficulty with this learning objective despite efforts to clarify during both workshops. Assignment scores and student opinion surveys show that this one-semester curriculum was an effective and productive means to introduce students to scientific literature and develop basic literacy skills.

Due to the complexity of research literature, this one semester implementation serves only as an introduction. We envision a continuation of the literature curriculum into the

second semester honors general chemistry course to address more advanced skills, such as correlating research questions to investigation methods, justifying how the experimental results address the hypotheses, evaluating presented figures and data tables, and comparing student conclusions from the presented results to the author's conclusions. With this scaffolded twosemester training, the undergraduates would be adequately equipped to continue to hone and apply these scientific literature evaluation skills in their future coursework, undergraduate research, and careers.

6.8 Acknowledgement

This work was supported in part by a grant from The University of Pittsburgh Discipline Based-Science Education Research Center (db-SERC). The authors thank the undergraduate teaching assistants for their help and support while implementing and assessing this curriculum.

7.0 Conclusion and Future Work

7.1 EDTA Carboxylate Symmetric Stretch

In chapter 4, the examination of unidentified peaks in the apo–EDTA, $[Ca:EDTA]^{2-}$, and $[Mg:EDTA]^{2-}$ carboxylate symmetric stretch regions elucidate their origin and connection to the metal binding geometry. We use FTIR and 2D-IR spectroscopy, implicit-solvent DFT analysis, ALMO–EDA theory, and introduce participation coefficient analysis to understand the carboxylate symmetric stretch vibrational bands in either the free or complexed state. We have developed an approach to provide participation coefficients analysis that allows us to untangle the mixing between the CH bending vibrations and the carboxylate symmetric stretches. In a way, this is similar to applying a "local-mode" picture to resolve the identity of the normal modes by mapping it onto the basis of the vibrations of each individual acetate arm of EDTA. The participation coefficients capture the subtle structural changes between Ca^{2+} and Mg^{2+} bound forms of EDTA.

Through the analysis of the 1000–1650 cm^{-1} region of the IR spectrum, we confirmed some of the previous assignments of the vibrational band and were able to assign the CH_t, CH_w, and CH_s modes and the mixed COO_s^- peaks. Using participation coefficients and DFT calculations, we show that the CH_s modes are mixed with the COO_s^-, whereas the COO_a^- modes are unmixed. 2D-IR spectroscopy with the participation coefficient analysis together can unravel the observed complexities in the carboxylate spectrum. Chapter 4 shows that coupling between the mixed modes create both diagonal and cross–peaks in the 2D-IR spectrum, whose positions are sensitive to the size of metal binding to EDTA. ALMO-EDA calculations also emphasize the primary importance of electrostatic interactions in the binding energy, while other non-electrostatic interactions are smaller but, nevertheless, dictate the depth of the ion in the EDTA binding pocket. Finally, this research revealed that the complexity of this mixed mode manifold provides the sensitivity needed to discern the changes in the EDTA structure when complexing with different cations (Ca^{2+} and Mg^{2+}).

The EDTA–divalent metal binding geometry research, which is a combined experimental

and theoretical effort to assign the vibrational modes of EDTA in the carboxylate symmetric stretch region, will be relevant to the biological community studying EF–hand metal interactions. In the EDTA molecule, the vibrational frequencies can now be correlated to the metal binding geometry. Chapter 4 also demonstrated that the changes in the diagonal and off-diagonal peaks from one divalent ion to another are primarily a result of geometric distortions. For example, this investigation can be a starting point for assigning the bands in the symmetric stretching region of EF–hand–Ca^{2+} or EF–hand–Mg^{2+} binding pockets. The sensitivity of the carboxylate symmetric stretch and CH bending mixed region to metal–binding geometry could be utilized to probe EF–hand–ion selectivity and sensitivity, protein activity, and allostery. In general, research relating to the chelation effects of carboxylate groups and the structural changes upon protein–ion interactions may benefit from the approach developed on this model compound.

7.2 pH Jump in HPTS–EAF

Chapter 5 explores the effect of a pH-jump experiment by a photoacid dissolved in a PIL solution. Using $TR^{M}PS$ and kinetic modeling, we interrogated the movement of a proton in EAF. In order to initiate the proton transfer reaction in EAF, a pyranine (HPTS) dye was shown to be efficient and recyclable on microsecond timescales. Assignment of the HPTS peaks in a well studied system such as 1 M acetate solution helped us identify the HPTS–EAF peaks. As expected from the solvent studies conducted with HPTS,[108] there were solvatochomic shifts of no more than 3–5 cm^{-1}in the HPTS, $HPTS^{*}$, PTS^{*-}, and PTS^{-} peaks in the EAF solution.

Time-resolved experiments were carried out to understand the mechanism of proton transfer in a PIL. In our data, we see that the HPTS ground state bleach and the $HPTS^{*}$ excited state absorption grows with an instrument response. Next we see the appearance of the PTS^{*-} species in a few picoseconds, which is followed by the appearance of the PTS^{-}. Our results also show that the formic acid species grows on an ultrafast ps timescale. We also observed the broad formic acid band changing into two separate peaks within a few

nanoseconds. These results indicate a slightly complex pathway of proton conduction through the EAF solution than depicted by our simple cyclic kinetic model. We also show that solvent dynamics is affecting the accurate extraction of kinetics from the current data.

The ionic nature of PILs facilitates the donor and acceptor framework, which is critical to the proton transfer process. Time-resolved two dimensional plots reveal the existence of all five different species in our reaction from timescales of a couple of picoseconds to around nanoseconds. This, along with the average kinetic fit suggests the existence of ultrafast proton transfer reaction (1 ps), excited state relaxation in 100 ps, and slow solvent equilibration kinetics in the EAF–HPTS system. Our $TR^{M}PS$ experiments indicate that the proton transfer process in EAF does not completely fit to a simple model of chemical kinetics, as it does not account for the different HCOOH peaks. New measurements on the fluorescence lifetime would provide new constraints on the rates of decay of $HPTS^{*}$ and PTS^{*-}. Our current analysis also shows that solvent contribution to the kinetics needs to be properly disentangled to extract accurate trends and to achieve that in the future different multivariate (SVD, PCA, Matrix factorization) data analysis approaches might help. Properly incorporating these processes may provide the accurate rates needed to improve the modeling of our observed kinetics. Alternatively, a kinetic scheme that explicitly treats the different types of 'contact pairs' and 'tight' or 'loose' complexes participating in the proton transfer in solution might be applicable. In aqueous proton transfer literature, the solvent separated species might form 'solvent bridges' or 'solvent wires'. Carrying out *ab initio* MD simulations might also support the experimental kinetics elucidation.[227,272,273]

As a future extension of this pH–jump investigation in EAF, alkyl chain length effects on the proton transfer process will illuminate the contribution from the non–polar domains in proton conduction process. Existence of nanostructure domains is already well established in room temperature ionic liquids as well as in protic ionic liquids.[213,274,275] Investigation using alkyl chain length and anions of varying sizes and charge distributions will illustrate the role of nanostructure organization of PILs in the proton transfer pathways.

7.3 Undergraduate Curriculum Development

The main objective of this curriculum development and research was to teach the undergraduates how to appreciate scientific literature. In chapter 6, through our curriculum, students get prepared for reading and understanding the differences and similarities between secondary and primary scientific literature. Chapter 6 also shows how undergraduates are equipped with the skills to pick out hypotheses, main goals, conclusions, and future directions of an original research publication. Our assessments, and semester and workshop surveys reveal that the students develop critical thinking skills through correlating hypotheses and conclusions. Students also formulate conclusions about the primary research after reading the publication. Students in the HGC course have appreciated this curriculum project and the skills they have developed through it. The undergraduates stated their improvement in research and critical thinking skills both during this curriculum implementation and during their active work at Pitt research laboratories after undergoing the training. This research is the start of a deeper scaffolded curriculum to teach students how to extrapolate ideas from scientific research.

Some of the challenges of this integrated curriculum are as follows: (1) the instructor must prepare to tackle the student confusion regarding the many complexities of scientific results. Different avenues such as writing order, result interpretation connections between results, conclusion, and hypothesis are new for first-time readers of an article. (2) Undergraduates must get trained to cope with the variety of writing styles, which can either be ideal or far from that. For example, most research papers do not have the hypothesis statement directly stated; rather, it is expressed throughout several sentences near the end of the introduction. This requires that the students must gain the skills to accurately interpret the prose and purpose of those paragraphs. (3) The students also should understand the relevance of general concepts not discussed in the paper underlying a testable scientific hypothesis and how those complicated concepts are addressed in the paper.

In the second semester, students will read the same articles from the first semester and focus on specific details of the research and critically evaluate them. Using the same articles from the first semester provides familiarity and some repetition. Therefore, students will be

able to focus on and apply more complicated analyses of the literature without becoming overwhelmed with the assignments.

Overall, this scaffolded teaching method, when applied to the course consistently, is expected to show the students are prepared to read and extrapolate from original research. The students will be able to understand, connect, and state hypotheses, methods, results, and conclusions. The integrated two-semester curriculum would continue to be a critical part of the HGC course training at the Pitt Chemistry department. Finally, this approach is new and unique in a sense that it gradually introduces students early on at the freshman level to the nuances of secondary and original scientific research and connects the relevance of that research to the world. This curriculum adds to the various initiatives in the literature that assist undergraduates in developing their scientific evaluation skills. As a future work, we can ask Pitt alumni students who have gone through this curriculum to state how this curriculum helped them to get situated in a critical problem-solving and research role.

7.4 Summary

Through this dissertation, we have explored the ion–binding structure in model compounds, the kinetics of proton transfer in a PIL, and developed a curriculum to educate undergraduates on evaluating primary research. Our first two scientific questions involve the use of time-resolved ultrafast spectroscopic tools to address the questions. We have implemented a participation coefficient analysis and explained the carboxylate symmetric stretch vibrational spectrum of EDTA, with 2D-IR spectroscopy and DFT calculations. With the $TR^{M}PS$ method, we have gained insights into the proton transfer kinetics in EAF, which indicated that the proton can travel through complicated pathways inside the protic ionic liquids. Finally, the curriculum developed in this research is a starting point for HGC students to get used to reading and understanding scientific literature.

Appendix

A.1 CH Mode Mixing Determines the Bandshape of Carboxylate Symmetric Stretch in Apo, Ca^{2+}, & Mg^{2+}–EDTA

A.1.1 Energy Decomposition Analysis

Table A.1 compares monomer- and dimer-basis SAPT0 results. Monomer-basis SAPT was originally chosen for presentation as it is closer to the ALMO-EDA restriction of the polarization term not using basis functions on other fragments. However, there is no qualitative difference between the monomer- and dimer-basis results. Additionally, the small values of $\Delta E_{\mathrm{CT}}^{\mathrm{SAPT}}$ indicate that the def2-TZVP basis set is sufficient to describe each monomer's wavefunction.

A.1.2 Participation Coefficient Analysis

First we show the COO average plot (Figures A.1), including the contributions of all the carboxylates together which gives an overall effect of the individual carboxylates from the plot in main text of chapter 4.

Figures A.2, A.3, A.4, and A.5 show that participation coefficient analysis from the ωB97X-D functional is in overall agreement with the M05-2X results from the paper. In the carboxylate stretch plots, the slight differences between the two methods are from the delocalization of the antisymmetric and symmetric COO^- modes. The COO_a^- modes in the ωB97X-D calculation have delocalization between all four carboxylates in modes 26–29. In the COO_s^- modes, the minor delocalization between the individual carboxylates is different from M05-2X. In effect, both M05-2X and ωB97X-D functionals assign the same mode numbers as the significant frequencies contributing to the COO_s^- motion.

Slightly different normal modes from M05-2X and ωB97X-D leads to slightly different frequencies (Tables A.3 and A.4), resulting in the peak assignments of CH_t, CH_w, and the CH_s modes (Figures A.4 and A.5) changing by one mode number in the participation

Table A.1: Comparison of monomer- and dimer-basis SAPT0 results, with definition of terms as in Table 4.4.

contribution	monomer basis		dimer basis	
(kcal mol^{-1})	Mg^{2+}	Ca^{2+}	Mg^{2+}	Ca^{2+}
ΔE_{geom}	136.33	96.51	134.15	94.20
ΔE_{solv}	0.00	0.00	0.00	0.00
ΔE_{elec}	−914.31	−862.97	−905.99	−856.66
ΔE_{Pauli}	68.43	102.75	64.02	96.48
ΔE_{disp}	−3.14	−12.13	−3.80	−13.39
ΔE_{pol}	−189.00	−143.14	−190.55	−140.61
ΔE_{CT}	−5.22	−3.47	−5.22	−3.47
ΔE_{int}	−901.67	−818.99	−902.16	−819.98

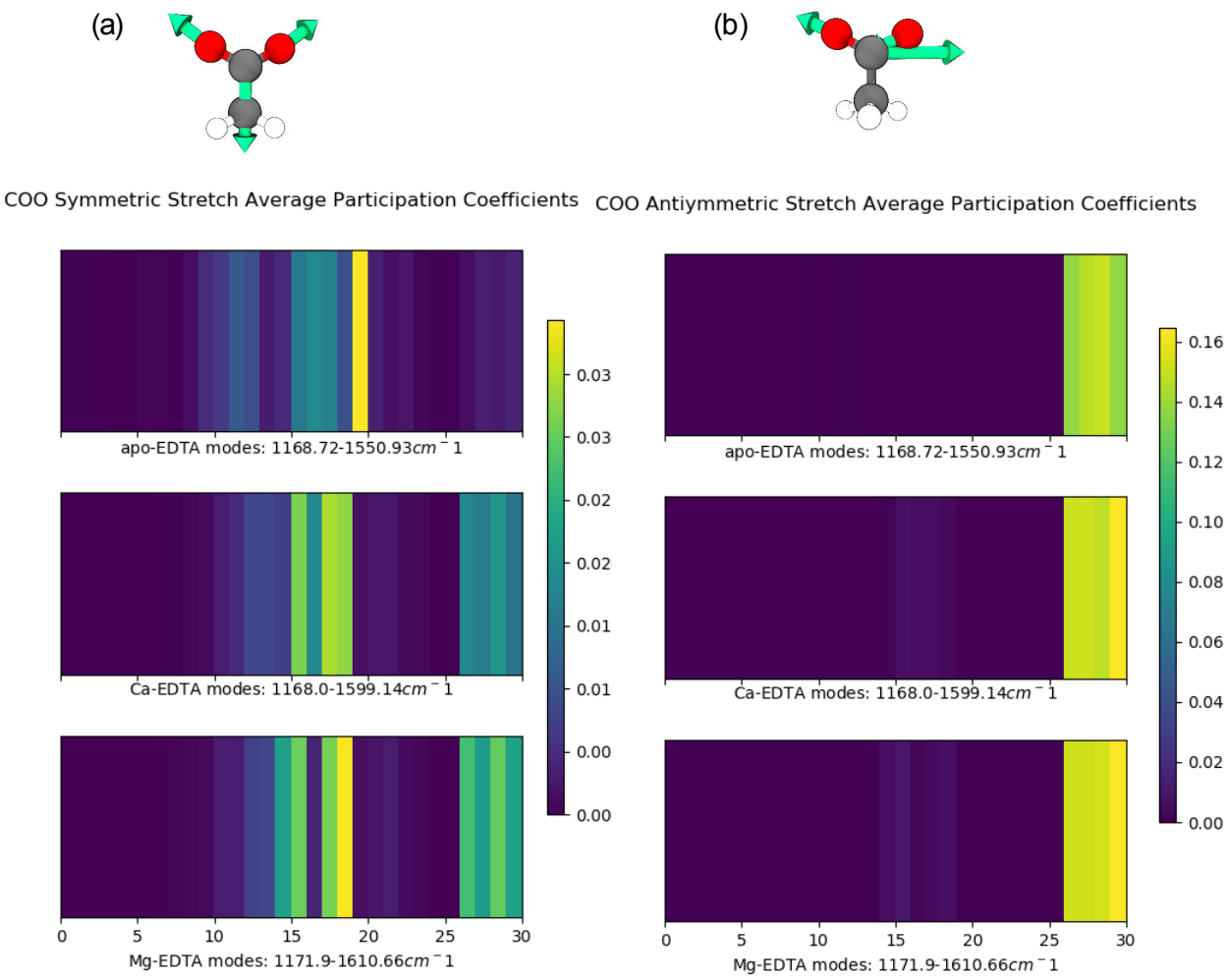

Figure A.1: M05-2X COO average participation coefficient shows more than five modes participate in the carboxylate symmetric stretch vibration. After metal binding, the four modes take up the majority of the COO_s^-.

coefficient plots. With ωB97X-D, both apo-EDTA and $[Mg{:}EDTA]^{2-}$ lose wag motions in their carboxylate symmetric region compared to M05-2X. However, the overall peak assignment and interpretation of the participation analysis (Figure A.6) remains the same. With both functionals, we find the vibration dominantly mixed with the carboxylate symmetric stretch is the CH_s mode and that the CH_t mode is sensitive to the identity of the bound metal (Figure A.5).

A participation analysis after the metal ion elimination from the optimized geometry frequency calculations (M05-2X and ωB97X-D) shows that the participation coefficients are sensitive to only the binding geometry and not the metal ion's identity (Figures A.7–A.12). The features in the individual participation plots are identical to the pattern of the participation coefficients calculated in the main text.

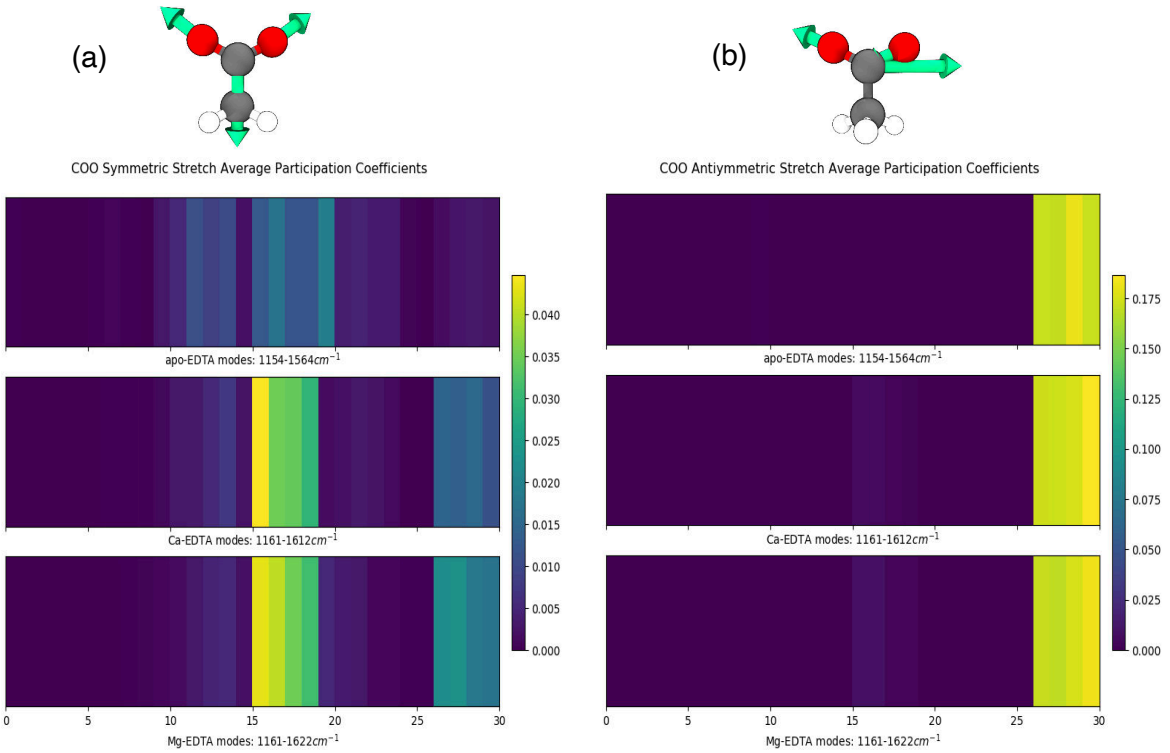

Figure A.2: ωB97X-D COO average participation coefficient shows more than five modes participate in the carboxylate symmetric stretch vibration. After metal binding, the four modes take up the majority of the COO^-_s.

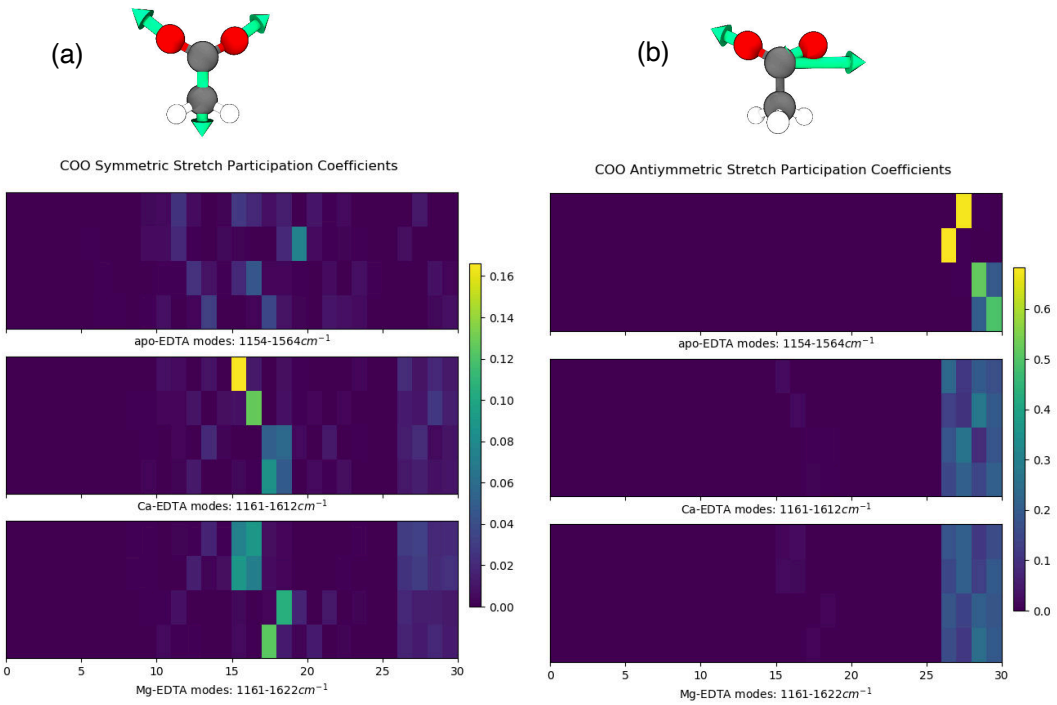

Figure A.3: ωB97X-D COO participation coefficients

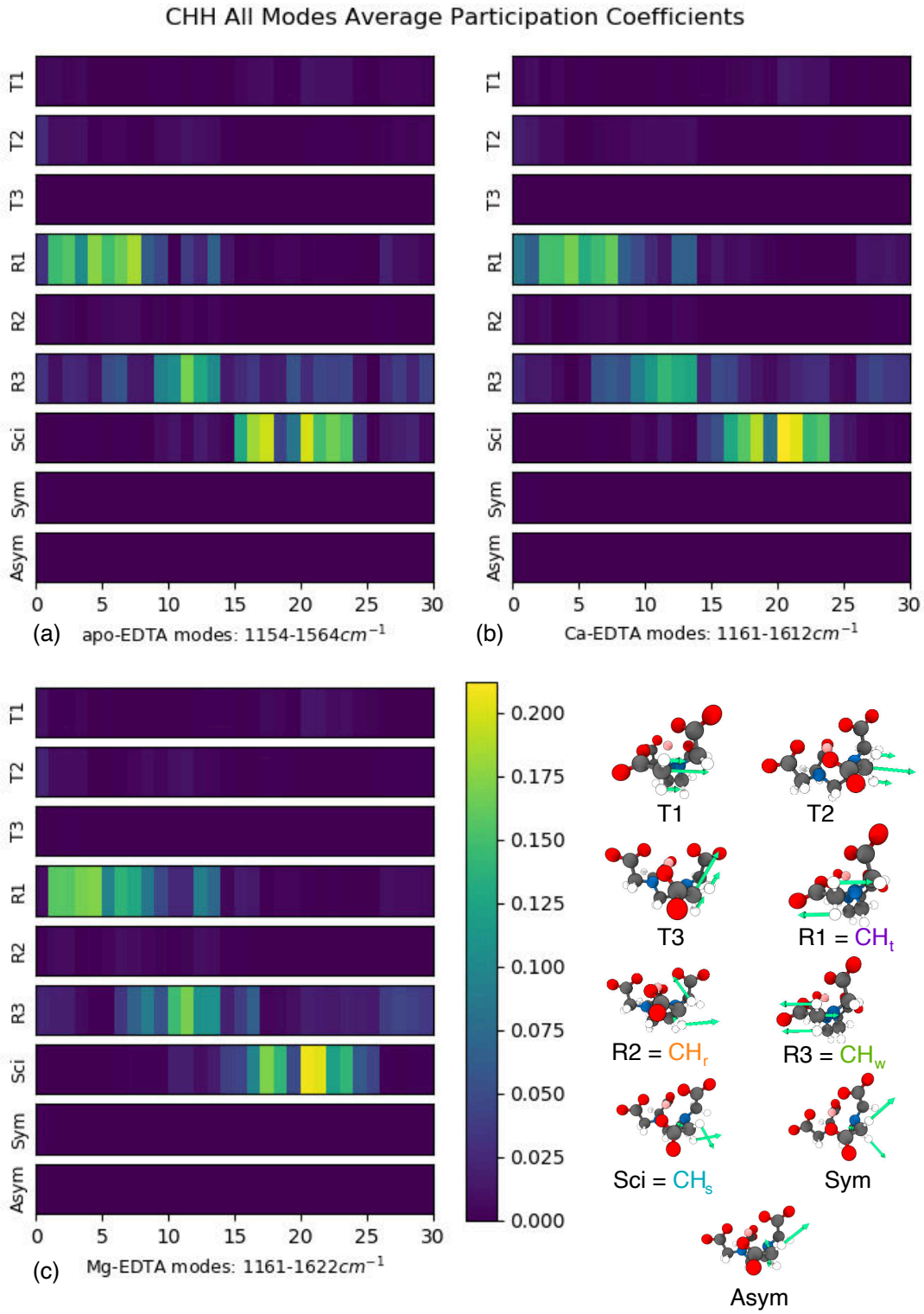

Figure A.4: ωB97X-D CHH average participation coefficients

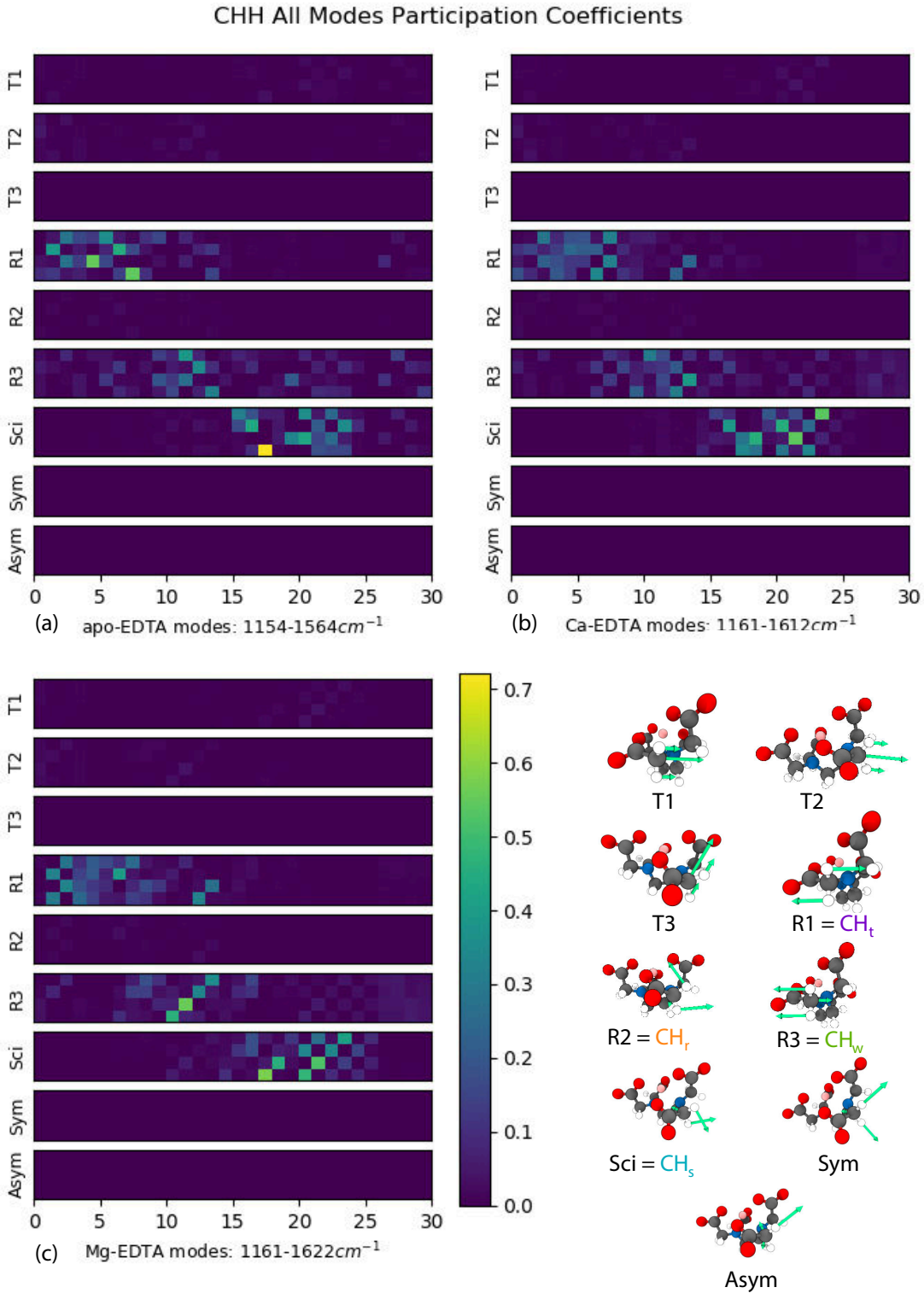

Figure A.5: ωB97X-D CHH participation coefficients

Table A.2: The largest M05-2X % participation coefficients (% pc) for COO_s, CH_s and CH_w in apo-EDTA, $[Ca:EDTA]^{2-}$, $[Mg:EDTA]^{2-}$ molecules.

normal mode index	Unscaled Mode (cm^{-1})	COO_s % pc	CH_s % pc	CH_w % pc
$[apo\text{-}EDTA]^{4-}$				
15	1446	6.0	32.7	15.8
16	1452	5.0	–	–
17	1456	6.4	62.8	14.7
19	1464	15.1	–	1.9
26	1546	54.7	–	–
27	1547	58.2	10.7	12.7
28	1549	60.1	–	–
29	1551	54.8	12.6	15.8
$[Ca:EDTA]^{2-}$				
15	1434	7.0	3.6	6.5
16	1440	3.9	17.2	12.9
17	1442	8.5	33.8	5.6
18	1445	7.7	42.8	4.6
26	1580	24.7	–	7.1
27	1584	36.2	–	13.6
28	1585	30.8	–	11.3
29	1599	17.5	–	5.3
$[Mg:EDTA]^{2-}$				
14	1418	3.8	–	2.2
15	1419	6.3	7.7	10.2
17	1437	9.8	53.6	–
18	1438	12.7	45.5	–
26	1592	18.6	–	6.2
27	1598	28.6	–	12.1
28	1600	25.9	–	8.6
29	1611	18.1	–	5.2

Table A.3: (Un)scaled M05-2X normal modes used in the participation coefficient (pc) calculations. The COO_s^- modes for (1) apo-EDTA = x-axis 9-13; (2) $[Ca:EDTA]^{2-}$ = x-axis 15-18; (3) $[Mg:EDTA]^{2-}$ = x-axis 14, 15, 17, 18. All the calculated COO_a^- modes are from 26-29 in the three molecules in the cm^{-1} units. The scaling factor is 0.9825.

normal mode	apo-EDTA		$[Ca:EDTA]^{2-}$		$[Mg:EDTA]^{2-}$	
index	unscaled	scaled	unscaled	scaled	unscaled	scaled
0	1169	1148	1168	1148	1172	1151
1	1178	1157	1173	1152	1174	1153
2	1194	1173	1176	1155	1175	1154
3	1226	1205	1230	1208	1235	1213
4	1281	1259	1267	1245	1263	1241
5	1302	1280	1299	1276	1294	1271
6	1310	1287	1311	1288	1318	1295
7	1314	1291	1317	1294	1320	1297
8	1346	1323	1348	1325	1338	1315
9	1366	1342	1363	1339	1361	1337
10	1370	1346	1367	1343	1364	1340
11	1371	1347	1376	1352	1373	1349
12	1377	1353	1382	1358	1381	1357
13	1392	1367	1385	1360	1383	1359
14	1399	1374	1426	1401	1418	1393
15	1446	1421	1434	1409	1419	1394
16	1452	1427	1440	1414	1425	1400
17	1456	1430	1442	1417	1437	1412
18	1460	1435	1445	1420	1438	1413
19	1464	1438	1456	1431	1446	1421
20	1475	1449	1476	1451	1478	1452
21	1479	1453	1480	1454	1479	1453
22	1485	1459	1490	1464	1490	1463
23	1497	1471	1492	1466	1493	1467
24	1499	1472	1520	1493	1521	1495
25	1511	1485	1526	1499	1528	1502
26	1546	1519	1580	1552	1592	1564
27	1547	1520	1584	1556	1598	1570
28	1549	1521	1585	1557	1600	1572
29	1551	1524	1599	1571	1611	1582

Table A.4: (Un)scaled ωB97X-D normal modes used in the participation coefficient (pc) calculations. The COO_s^- modes for (1) apo-EDTA = x-axis 9-13; (2) $[Ca:EDTA]^{2-}$ = x-axis 15-18; (3) $[Mg:EDTA]^{2-}$ = x-axis 14, 15, 17, 18. All the calculated COO_a^- modes are from 26-29 in the three molecules. All values have units of cm^{-1}. The scaling factor is 0.9927.

normal mode	apo-EDTA		$[Ca:EDTA]^{2-}$		$[Mg:EDTA]^{2-}$	
index	unscaled	scaled	unscaled	scaled	unscaled	scaled
0	1154	1146	1161	1153	1161	1153
1	1179	1170	1164	1156	1164	1156
2	1182	1173	1169	1160	1167	1158
3	1230	1221	1222	1213	1223	1214
4	1278	1268	1255	1246	1251	1242
5	1293	1284	1289	1280	1286	1277
6	1297	1288	1301	1292	1312	1302
7	1306	1296	1305	1295	1320	1310
8	1318	1308	1336	1326	1330	1320
9	1353	1343	1352	1342	1347	1337
10	1363	1353	1358	1348	1352	1342
11	1367	1357	1366	1356	1361	1351
12	1371	1361	1373	1363	1367	1357
13	1379	1369	1382	1372	1373	1363
14	1397	1387	1414	1404	1402	1392
15	1431	1421	1422	1412	1409	1399
16	1433	1423	1424	1414	1411	1401
17	1435	1425	1430	1420	1426	1416
18	1439	1428	1432	1422	1428	1418
19	1441	1430	1439	1428	1432	1422
20	1457	1446	1464	1453	1465	1454
21	1462	1451	1465	1454	1469	1458
22	1465	1454	1471	1460	1473	1462
23	1473	1462	1478	1467	1479	1468
24	1492	1481	1504	1493	1494	1483
25	1506	1495	1514	1503	1505	1494
26	1562	1551	1592	1580	1604	1592
27	1563	1552	1596	1584	1610	1598
28	1563	1552	1597	1585	1611	1599
29	1564	1553	1612	1600	1622	1610

The Mg^{2+} cation is closely packed in the complex (Table A.5). The O–Mg^{2+} distance is small (∼2.06 Å) whereas the O–Ca^{2+} distance is ∼2.34 Å. Also, the diametrically oppo-

site carboxylate C-C distances are shorter in case of $[Mg:EDTA]^{2-}$ (∼5.30 Å and 5.60 Å) compared to the $[Ca:EDTA]^{2-}$ (∼5.40 Å and 6.30 Å). There is a smaller spread of dihedral angles $\angle N-C-C-O$ in $[Mg:EDTA]^{2-}$ (∼ 18–22.50°) compared to the calcium cation ($\angle N-C-C-O$ = ∼24–29°). The N–Mg^{2+} distance is 0.29 Å shorter than the N–Ca^{2+} distance which reflects that the magnesium cation sits in a tighter space in compared to the calcium ion. Our dihedral angle outputs also show that the EDTA binding pocket is distorted more by the Mg^{2+} ion than the Ca^{2+} ion. The magnesium ion (72 pm effective radius[191]) can easily fit inside the tetra-carboxylate binding pocket of EDTA, however, the ∼38% size increase of Ca^{2+} (100 pm effective radius[191]) makes it difficult to pack the ion tightly inside this pocket.

A.1.3 FTIR Analysis

Table A.6 shows that the Gaussian best fits from experimental spectra. Figure A.14 shows those same best fit curves overlaid on the calculated normal modes from both functional. Both the functionals have the same trend and the assignments of modes remain the same.

Table A.5: Sagittal (S) and equatorial (E) dihedral angles between oxygens-nitrogens for the two complexes show geometry differences are comparable in both M05 and ωB97X-D methods.

	M05-2X		ωB97X-D	
	Ca^{2+}	Mg^{2+}	Ca^{2+}	Mg^{2+}
(S)∠NCCO	24.0°	21.7°	24.7°	22.9°
(S)∠NCCO	26.2°	22.3°	25.2°	23.9°
(E)∠NCCO	28.6°	18.2°	26.4°	15.1°
(E)∠NCCO	28.3°	18.1°	26.2°	14.0°
$\angle NM^{2+}N$	72.1°	80.6°	72.5°	80.1°
$N-M^{2+}$ (Å)	2.51	2.22	2.53	2.25
$O-M^{2+}$ (Å)	2.32, 2.34	2.05	2.32, 2.34	2.06
$C-M^{2+}$ (Å)	3.17	2.84, 2.87	3.17	2.86, 2.90

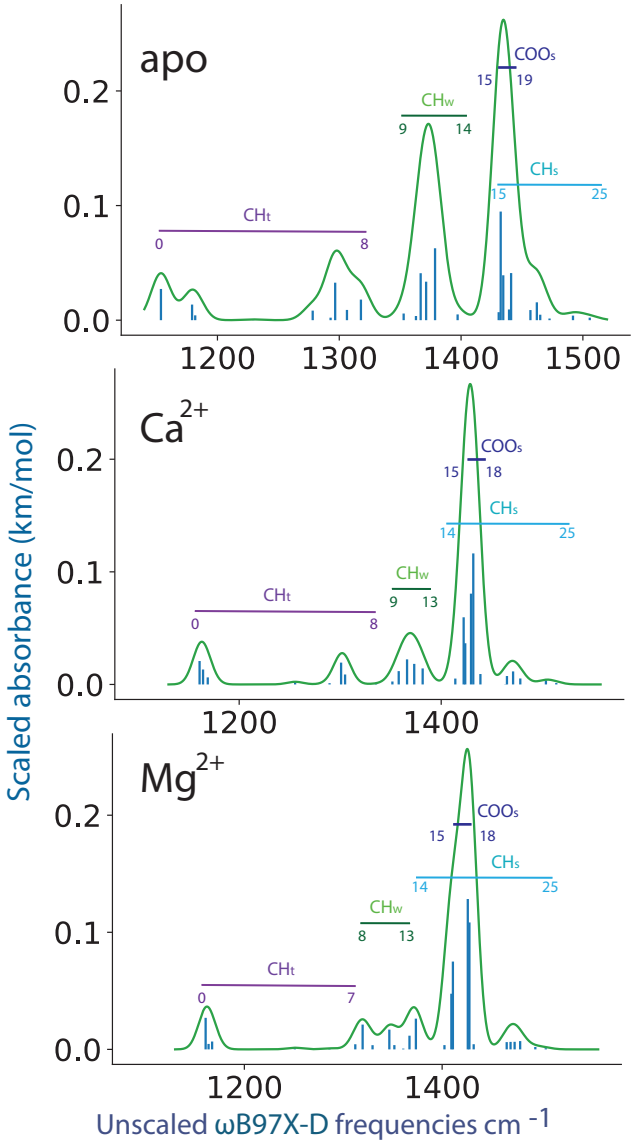

Figure A.6: ωB97X-D symmetric stretch region assignments

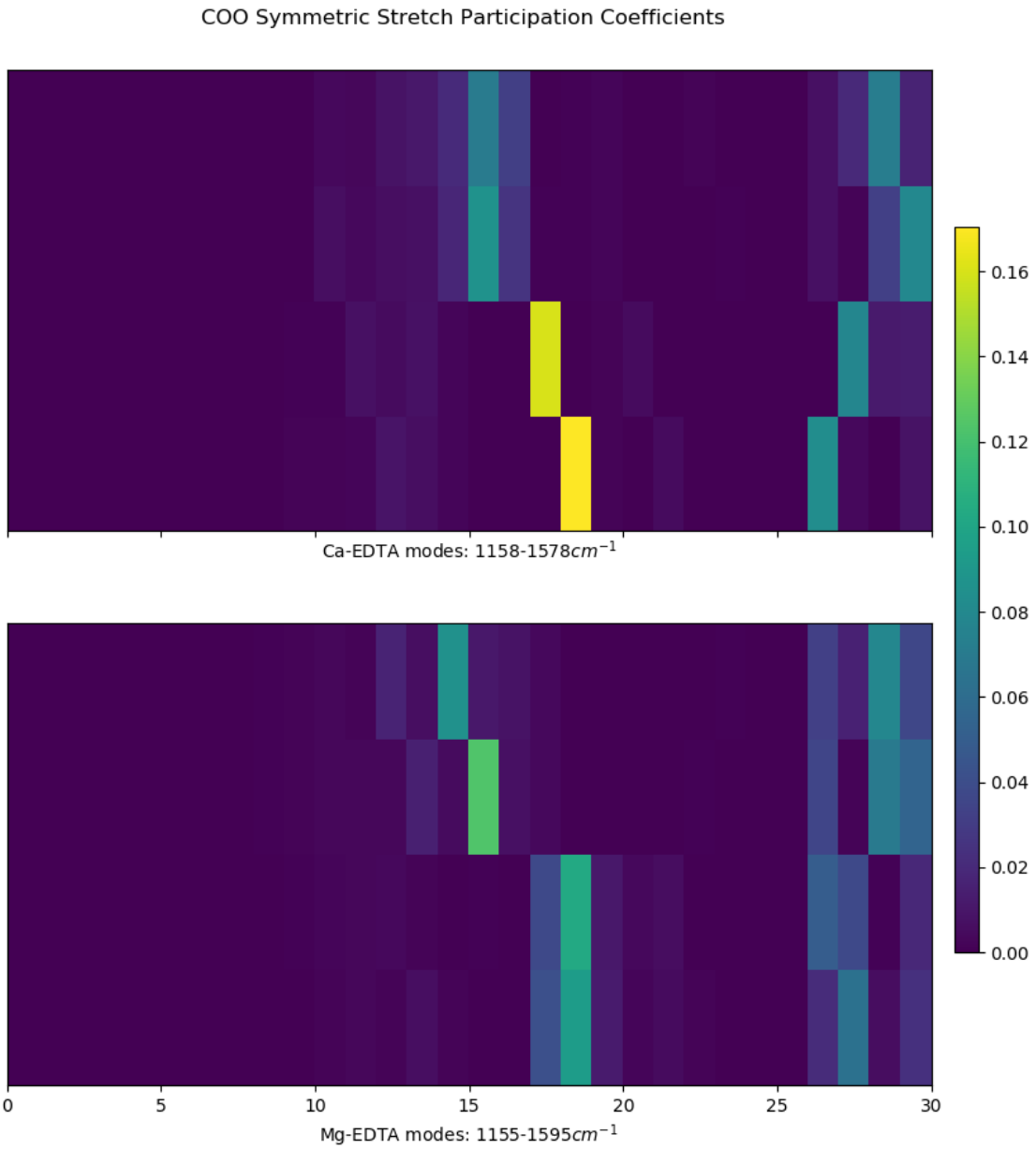

Figure A.7: Metal-less COO participation coefficients from M05-2X functional.

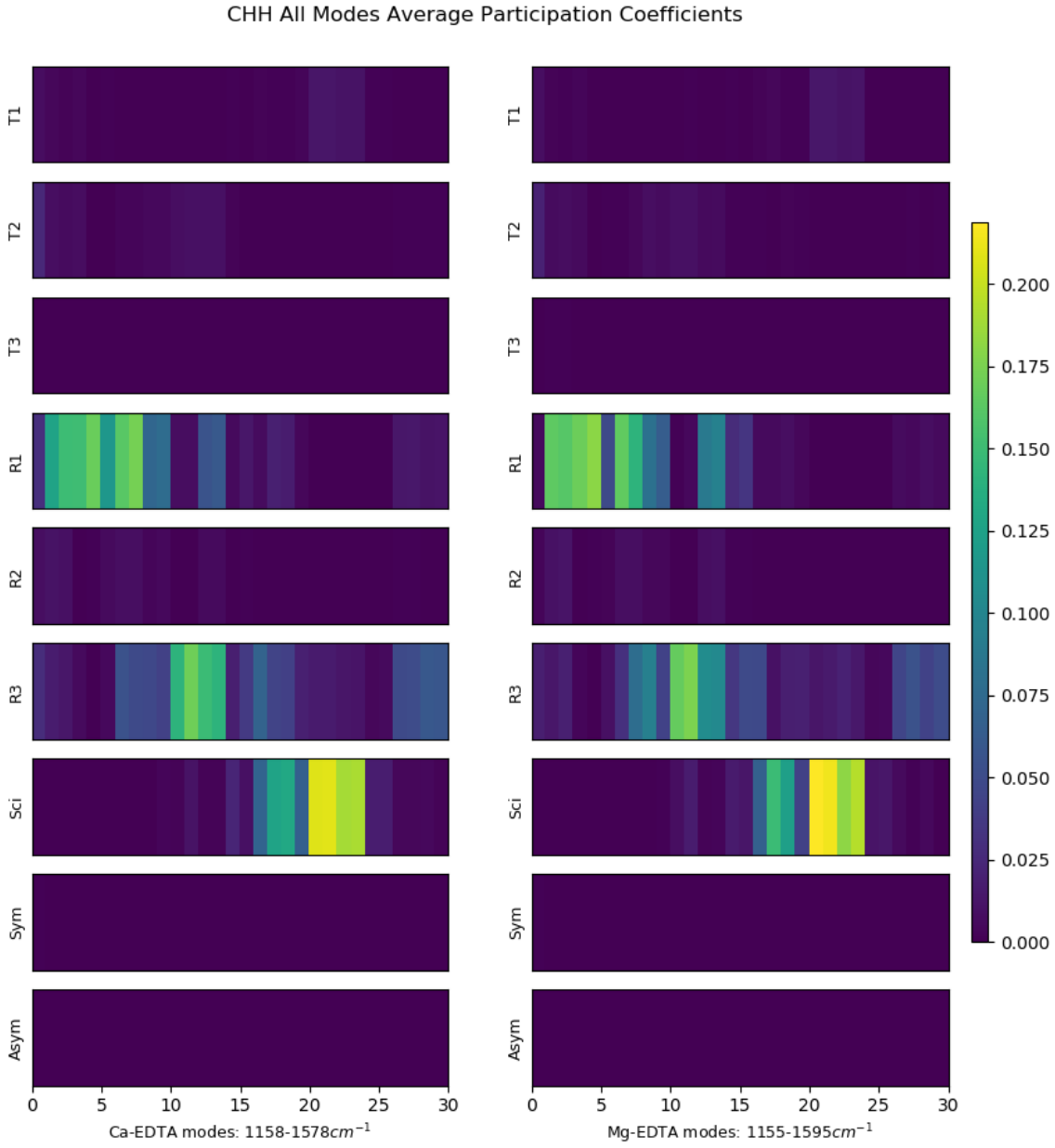

Figure A.8: Metal-less average CHH participation coefficients from the M05-2X calculations.

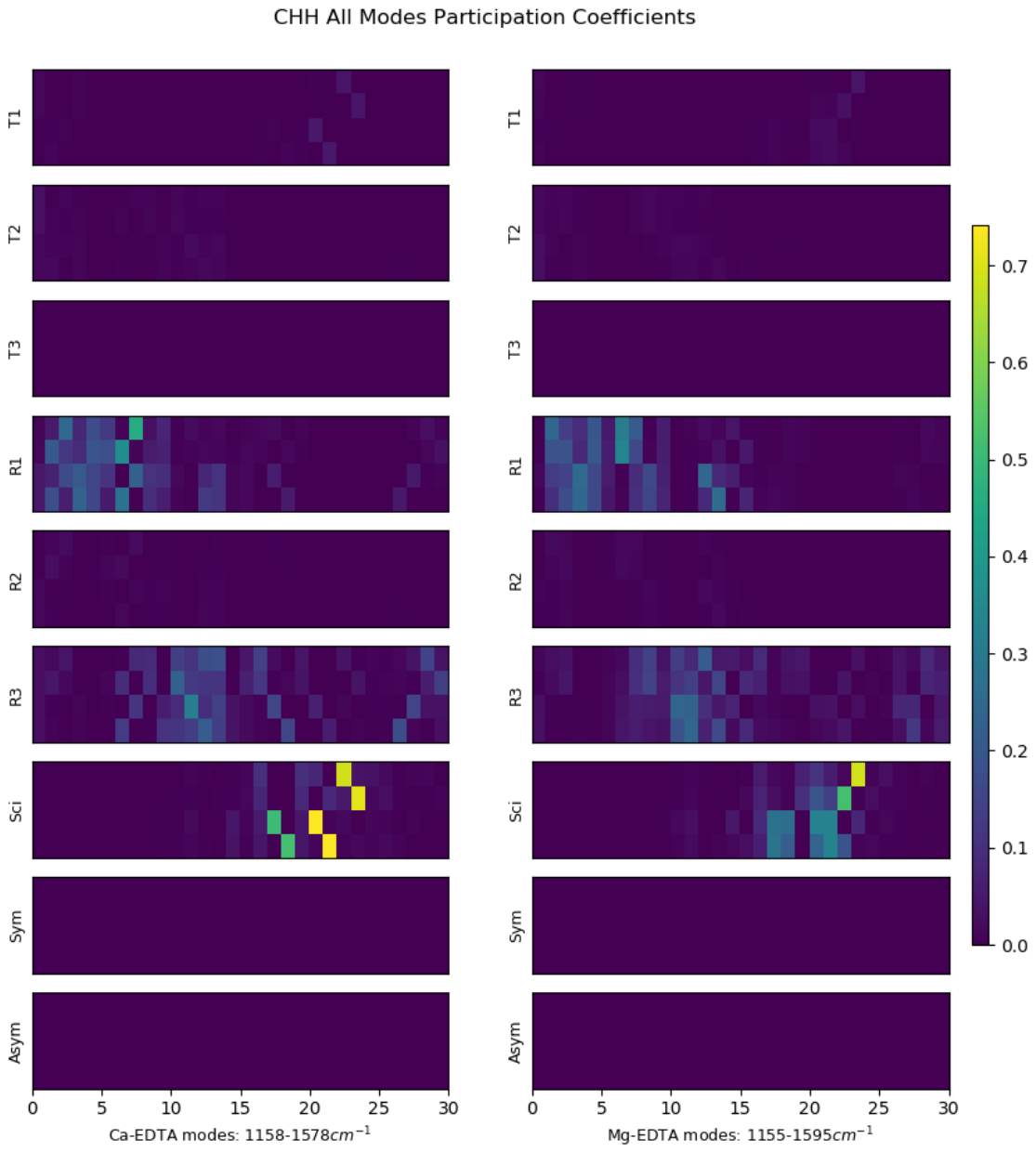

Figure A.9: Metal-less CHH participation coefficients from the M05-2X calculations.

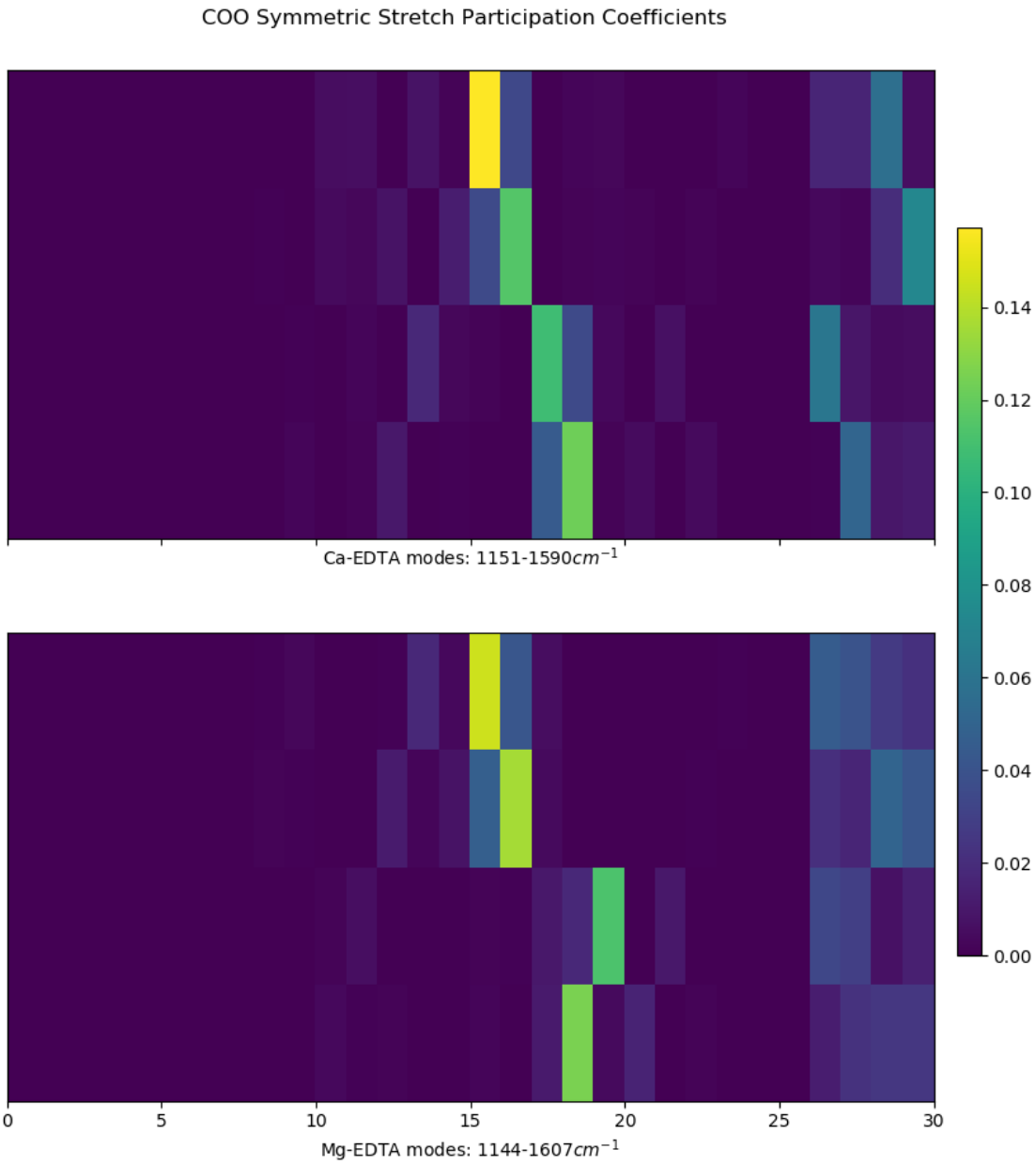

Figure A.10: Metal-less COO participation coefficients from ωB97X-D functional.

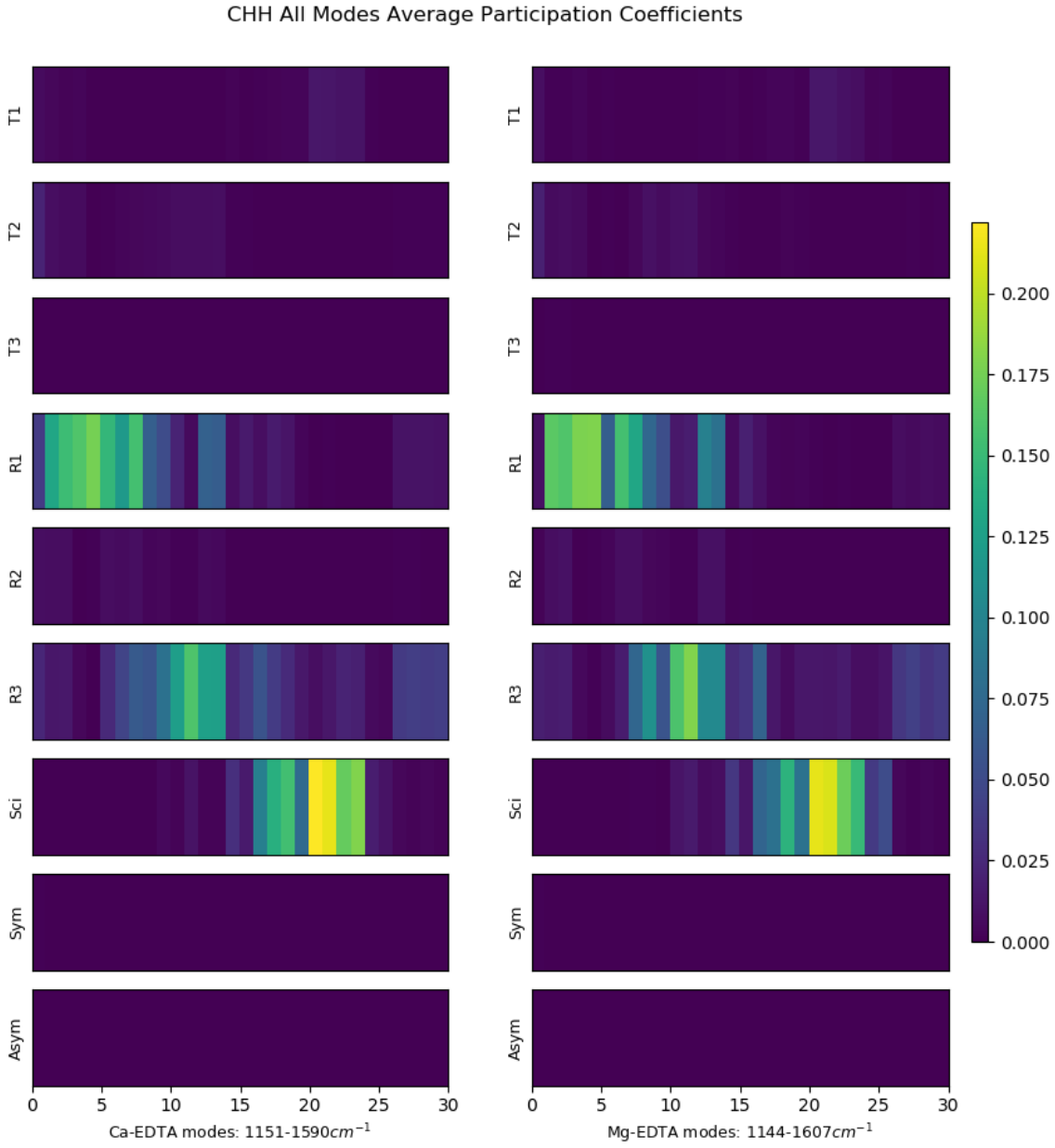

Figure A.11: Metal-less average CHH participation coefficients from ωB97X-D calculations.

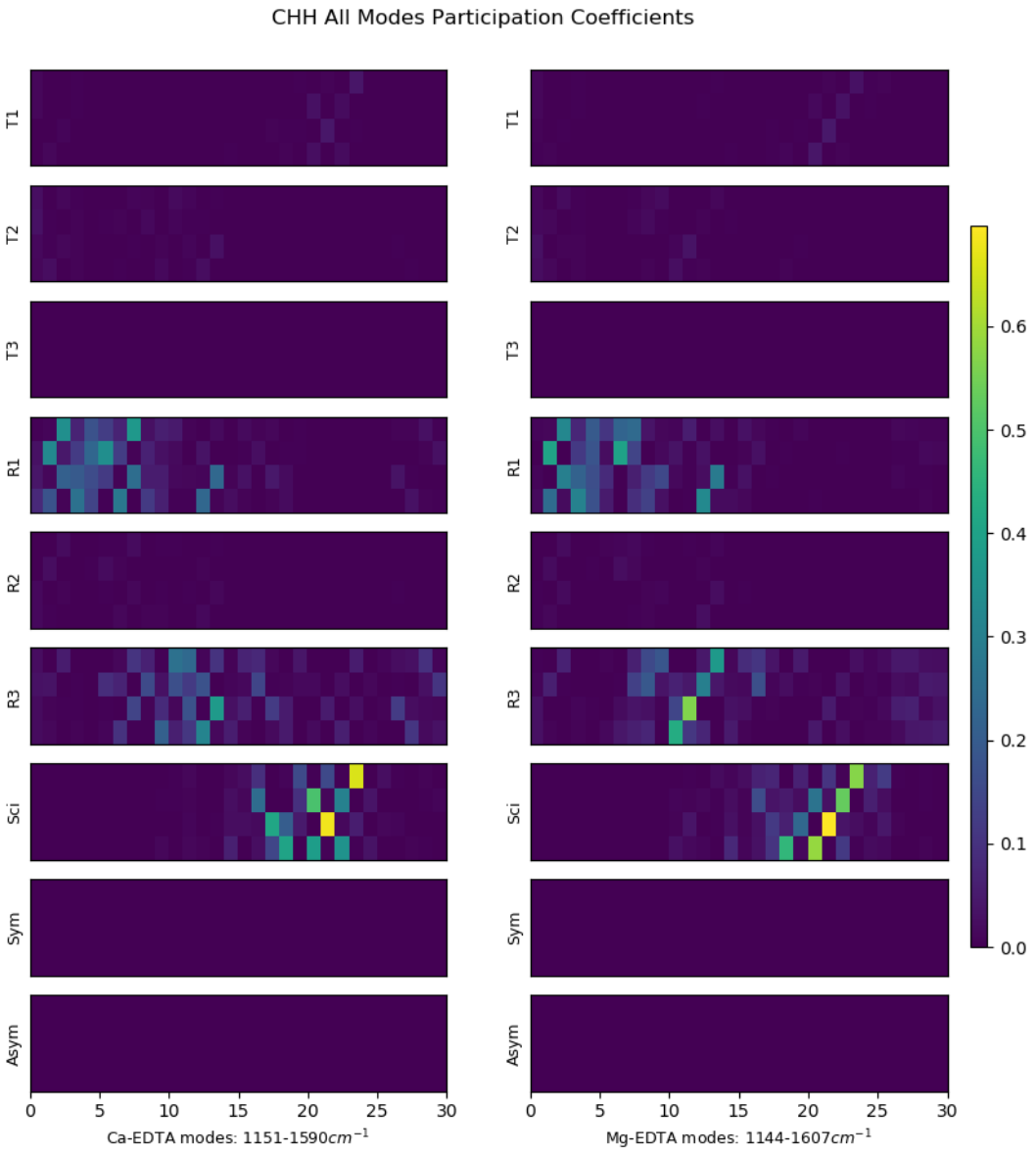

Figure A.12: Metal-less CHH participation coefficients from ωB97X-D calculations.

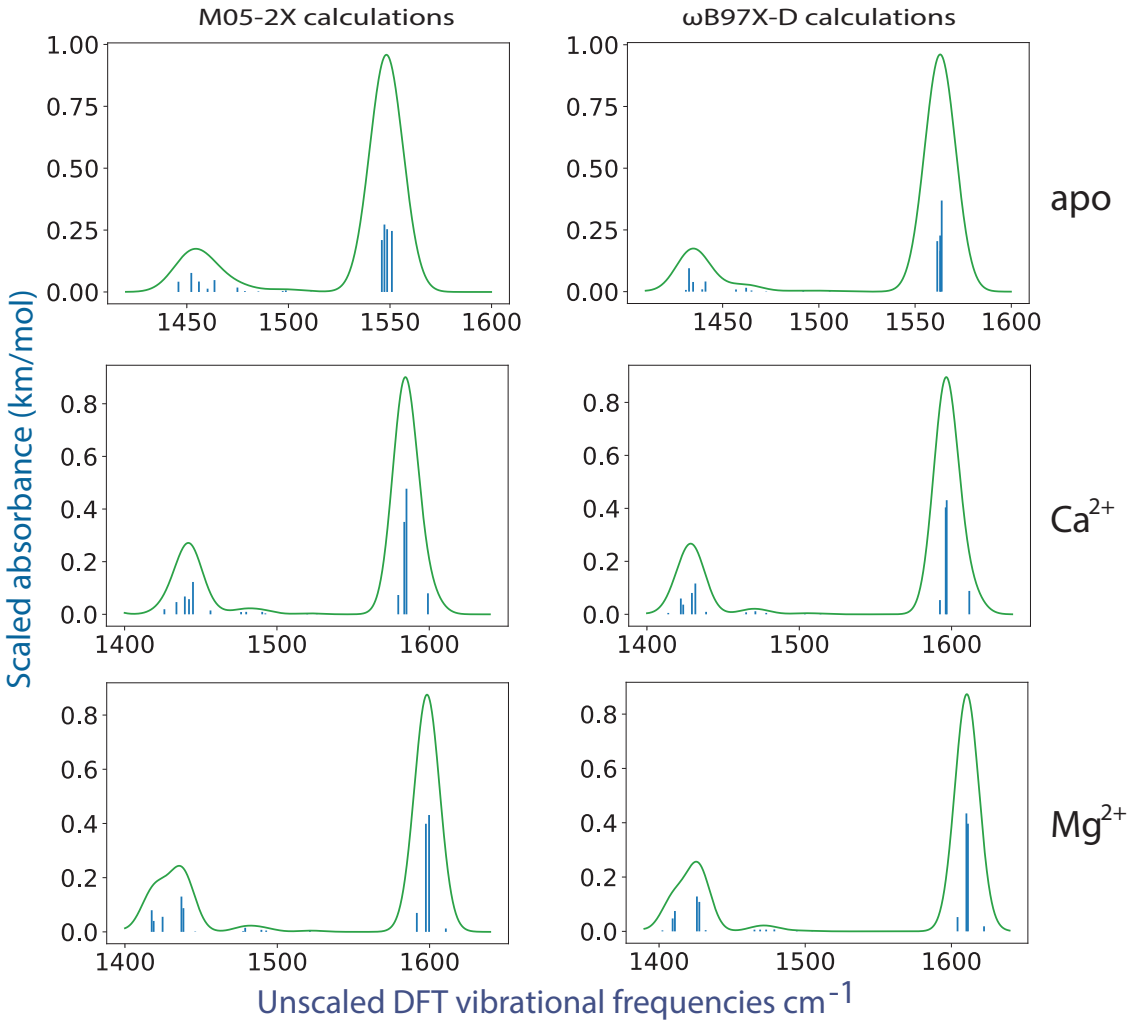

Figure A.13: Unscaled carboxylate region DFT frequencies of (apo) apo-EDTA, (Ca^{2+}) $[Ca:EDTA]^{2-}$, and (Mg^{2+}) $[Mg:EDTA]^{2-}$ (8 cm^{-1} Gaussian width, intensity scaling factor = 0.0002) in two different functionals have the same qualitative trend.

Table A.6: Center frequencies and full width at half maximum (FWHM) from Gaussian best fits in apo-EDTA, $[Ca:EDTA]^{2-}$, and $[Mg:EDTA]^{2-}$ for the symmetric stretch region in each molecule, in units of cm^{-1}.

center frequency	FWHM
$[apo\text{-}EDTA]^{4-}$	
1407	23.30
1429	11.63
1444	13.84
1457	28.26
$[Ca:EDTA]^{2-}$	
1414	19.97
1433	8.02
1442	10.92
1456	41.06
$[Mg:EDTA]^{2-}$	
1407	10.00
1418	23.31
1436	8.07
1447	12.23
1464	16.40

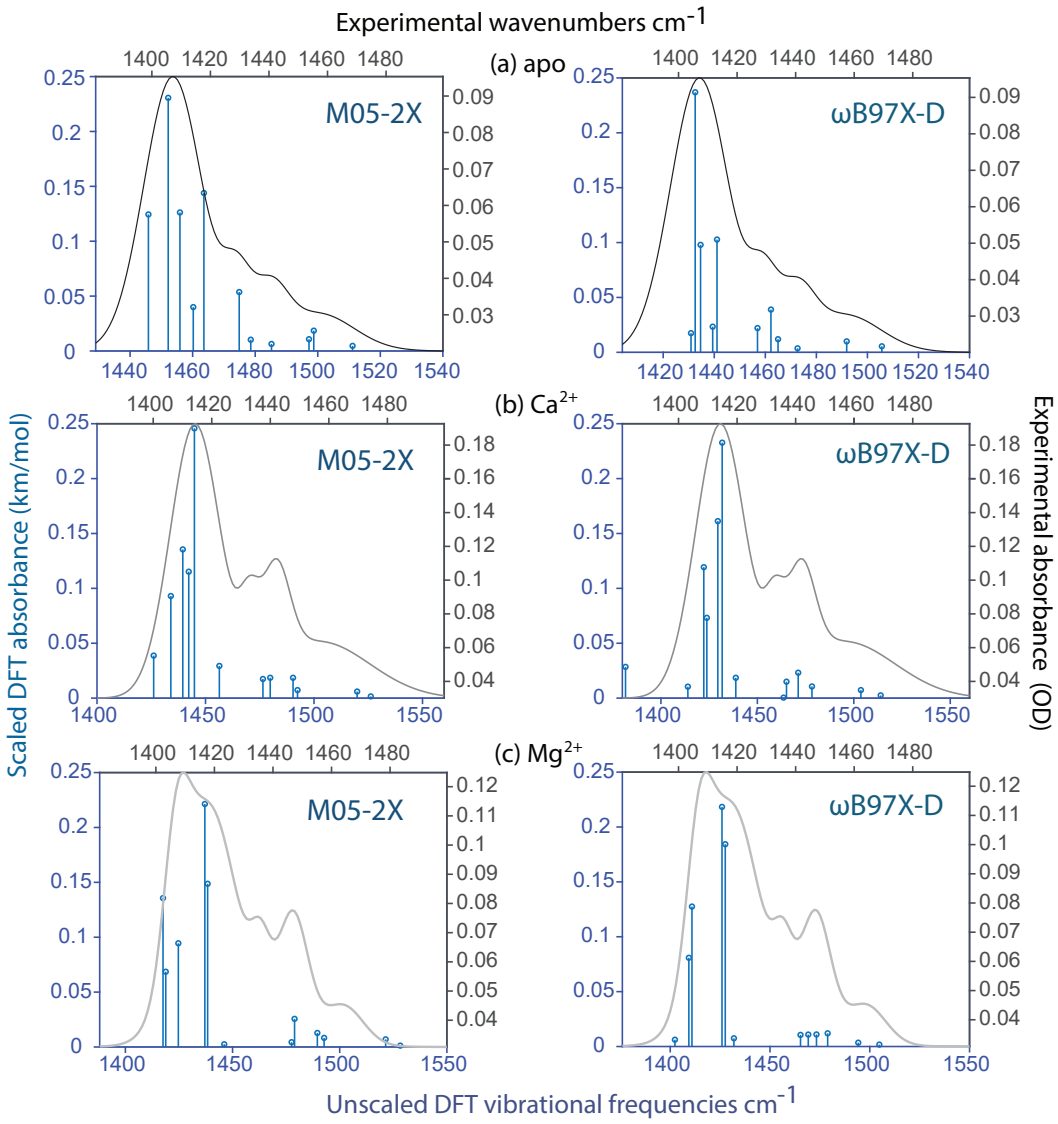

Figure A.14: Overlaid normal modes overlap significantly with the Gaussian best fit peak centers in all three molecules. The COO_s^- region of the experimental Gaussian best fits (8 cm^{-1}) are overlaid onto the unscaled calculated frequencies. The top row (a) is apo-EDTA, the middle row (b) is $[Ca:EDTA]^{2-}$ and the bottom row (c) is $[Mg:EDTA]^{2-}$. The left column shows the M05-2X/def2-TZVP/SMD theory results, while the right column contains results from the ωB97X-D/def2-TZVP/SMD calculations.

A.2 Calculation Files for Chapter 4

Appendix section A.2 contains the combined geometry optimization/frequency calculation inputs in the segments `apo_opt_freq_m.in`, `ca_opt_freq_m.in`, `mg_opt_freq_m.in`, `apo_opt_freq_w.in`, `ca_opt_freq_w.in`, and `mg_edta_w.in` for the apo-, Ca-, and Mg-EDTA molecules, respectively (`m` for M05-2X and `w` for ωB97X-D). The file used to establish spring and angle force coefficients for the submolecular basis calculations is `methane.in`. The EDA and SAPT calculations on the optimized complexes are `mg_eda_w.in`, `ca_eda_w.in`, `mg_sapt.in`, and `ca_sapt.in`, and the corresponding swapped metal calculations are `mginca_eda_w.in`, `cainmg_eda_w.in`, `mginca_sapt.in`, and `cainmg_sapt.in`.

```
-----------------------------------------------------
apo_opt_freq_m.in
-----------------------------------------------------
$molecule
-4 1
N        -0.0413476082   -0.0097195466   -0.0302620842
C        -0.0388987439    0.0555353972    1.4182459723
H         0.9766655645   -0.0976584627    1.7788906908
H        -0.3605482583    1.0455736393    1.7609271905
C         0.9414170791    0.9292799581   -0.5356910211
H         0.8779142597    0.9647213861   -1.6212699918
H         0.7428064705    1.9426126798   -0.1679013074
C        -1.3636863653    0.3925132072   -0.4972339960
H        -2.0887460737   -0.2549918151   -0.0145792824
H        -1.5725711172    1.4204353416   -0.1746035969
C        -0.8977744395   -0.9862775153    2.1228620657
O        -1.2173166836   -2.0395725942    1.5241188127
O        -1.1959253110   -0.7148005004    3.3161977307
C         2.3933309452    0.5903783465   -0.2253860124
O         2.7165534933   -0.5975406766    0.0087304198
O         3.1946057610    1.5603439967   -0.2758749860
C        -1.5090801488    0.2850006091   -2.0108714353
H        -1.1419756598    1.1919714782   -2.4833872638
H        -0.9034891866   -0.5462764692   -2.3650463503
N        -2.8791378174    0.0684935201   -2.4763316537
C        -3.3612058268   -1.2403084279   -2.0844283225
H        -2.5920066307   -1.9777936883   -2.3138524256
H        -3.5589799454   -1.3103692016   -1.0115144788
C        -3.7673219257    1.0791393529   -1.9274127190
H        -4.7705216677    0.9035768011   -2.3129795802
H        -3.8205757962    1.0156283600   -0.8387555482
C        -3.3965178548    2.5070456609   -2.3065431389
```

```
O        -3.0149206952    2.7390993387   -3.4792818219
O        -3.5443607931    3.3796697028   -1.4119679559
C        -4.6255418275   -1.7153651348   -2.7923304269
O        -5.0198124083   -1.1434544898   -3.8339573390
O        -5.1807701075   -2.7162174688   -2.2642598168
$end

$rem
 jobtype = opt
 method =  m05-2x
 basis = def2-tzvp
 solvent_method = smd
!purecart = 2111
 mem_static = 16000
 mem_total = 32000
 geom_opt_max_cycles = 600
 geom_opt_tol_gradient = 40
 geom_opt_tol_energy = 20
 geom_opt_tol_displacement = 100
 geom_opt_dmax = 10
 scf_convergence = 8
 max_scf_cycles = 10000
 xc_grid = 3
$end

$molecule
READ
$end

$rem
 jobtype = freq
 method =  m05-2x
 basis = def2-tzvp
 solvent_method = smd
 mem_static = 16000
 mem_total = 32000
 max_scf_cycles = 10000
 vibman_print = 6
 xc_grid = 3
$end

-------------------------------------------------------
 apo_opt_freq_w.in
-------------------------------------------------------
 $molecule
-4 1
 N        -0.0413476082   -0.0097195466   -0.0302620842
```

```
C        -0.0388987439    0.0555353972    1.4182459723
H         0.9766655645   -0.0976584627    1.7788906908
H        -0.3605482583    1.0455736393    1.7609271905
C         0.9414170791    0.9292799581   -0.5356910211
H         0.8779142597    0.9647213861   -1.6212699918
H         0.7428064705    1.9426126798   -0.1679013074
C        -1.3636863653    0.3925132072   -0.4972339960
H        -2.0887460737   -0.2549918151   -0.0145792824
H        -1.5725711172    1.4204353416   -0.1746035969
C        -0.8977744395   -0.9862775153    2.1228620657
O        -1.2173166836   -2.0395725942    1.5241188127
O        -1.1959253110   -0.7148005004    3.3161977307
C         2.3933309452    0.5903783465   -0.2253860124
O         2.7165534933   -0.5975406766    0.0087304198
O         3.1946057610    1.5603439967   -0.2758749860
C        -1.5090801488    0.2850006091   -2.0108714353
H        -1.1419756598    1.1919714782   -2.4833872638
H        -0.9034891866   -0.5462764692   -2.3650463503
N        -2.8791378174    0.0684935201   -2.4763316537
C        -3.3612058268   -1.2403084279   -2.0844283225
H        -2.5920066307   -1.9777936883   -2.3138524256
H        -3.5589799454   -1.3103692016   -1.0115144788
C        -3.7673219257    1.0791393529   -1.9274127190
H        -4.7705216677    0.9035768011   -2.3129795802
H        -3.8205757962    1.0156283600   -0.8387555482
C        -3.3965178548    2.5070456609   -2.3065431389
O        -3.0149206952    2.7390993387   -3.4792818219
O        -3.5443607931    3.3796697028   -1.4119679559
C        -4.6255418275   -1.7153651348   -2.7923304269
O        -5.0198124083   -1.1434544898   -3.8339573390
O        -5.1807701075   -2.7162174688   -2.2642598168
$end

$rem
 jobtype = opt
 method =  wB97X-D
 basis = def2-tzvp
 solvent_method = smd
 mem_static = 16000
 mem_total = 32000
 geom_opt_max_cycles = 600
 geom_opt_tol_gradient = 40
 geom_opt_tol_energy = 20
 geom_opt_tol_displacement = 100
 geom_opt_dmax = 10
 scf_convergence = 8
 max_scf_cycles = 10000
```

```
 xc_grid = 2
$end


$molecule
READ
$end

$rem
 jobtype = freq
 method =  wB97X-D
 basis = def2-tzvp
 solvent_method = smd
 mem_static = 16000
 mem_total = 32000
 max_scf_cycles = 10000
 vibman_print =   6
 xc_grid = 2
$end
```

ca_opt_freq_m.in

```
$molecule
-2 1
Ca       -0.0440310656    0.0125458763    0.0647396429
O         0.0484169175   -0.0750268653    2.4457312145
C         1.1617738866   -0.0493710059    3.0506870148
O         1.3144515429   -0.2414418733    4.2698740531
C         2.3961324965    0.3147917870    2.2326156243
H         3.2908418970   -0.0389744100    2.7460146534
H         2.4172258482    1.4045608095    2.2370944296
N         2.3796110769   -0.1342678605    0.8451813420
C         3.1333158137    0.7861703332    0.0110795110
H         3.2913257448    0.3271897909   -0.9637187368
H         4.1109404300    1.0228987328    0.4345610052
C         2.8657428534   -1.5075230320    0.7158359062
H         3.1275981479   -1.6776192089   -0.3243562660
H         3.7720633550   -1.6551719031    1.3083306617
C         1.8183477910   -2.5114655261    1.1447226403
H         1.5186925904   -2.3211364057    2.1708497481
H         2.2468312203   -3.5161811189    1.1055887585
N         0.6299237930   -2.4270430596    0.2988497905
C        -0.5068643211   -3.1089500045    0.8918051870
H        -0.5466360484   -2.8630623135    1.9520530538
H        -0.4343519661   -4.1928628665    0.7969670330
C         0.8935798983   -2.9386731454   -1.0412389843
```

```
H        -0.0486256857   -3.2782386446   -1.4728301653
H         1.5724219932   -3.7918065977   -1.0079488043
C         1.4233566415   -1.9087675700   -2.0319888411
O         2.2109505240   -2.3122028268   -2.9092736279
O         0.9825839671   -0.7268863108   -1.9453049545
C         2.3829641222    2.0860224616   -0.2594662391
O         3.0634394153    3.0701810270   -0.5997062929
O         1.1184803816    2.0773259083   -0.1668114645
C        -1.8331887285   -2.6464109009    0.2977970549
O        -2.7788672875   -3.4549326696    0.3165831130
O        -1.9058177974   -1.4584890016   -0.1384842446
$end


$rem
 jobtype = opt
 method =  m05-2x
 basis =  def2-tzvp
 solvent_method = smd
 mem_static = 1000
 mem_total = 16000
 geom_opt_max_cycles = 600
 geom_opt_tol_gradient = 40
 geom_opt_tol_energy = 20
 geom_opt_tol_displacement = 100
 geom_opt_dmax = 10
 scf_convergence = 8
 max_scf_cycles = 5000
 xc_grid = 3
$end


$molecule
 READ
$end


$rem
 jobtype = freq
 method =  m05-2x
 basis = def2-tzvp
 solvent_method = smd
 mem_static = 1000
 mem_total = 16000
 max_scf_cycles = 5000
 vibman_print =   6
 xc_grid = 3
```

```
$end
```

---

ca_opt_freq_w.in

```
$molecule
-2 1
Ca       -0.0440310656    0.0125458763    0.0647396429
O         0.0484169175   -0.0750268653    2.4457312145
C         1.1617738866   -0.0493710059    3.0506870148
O         1.3144515429   -0.2414418733    4.2698740531
C         2.3961324965    0.3147917870    2.2326156243
H         3.2908418970   -0.0389744100    2.7460146534
H         2.4172258482    1.4045608095    2.2370944296
N         2.3796110769   -0.1342678605    0.8451813420
C         3.1333158137    0.7861703332    0.0110795110
H         3.2913257448    0.3271897909   -0.9637187368
H         4.1109404300    1.0228987328    0.4345610052
C         2.8657428534   -1.5075230320    0.7158359062
H         3.1275981479   -1.6776192089   -0.3243562660
H         3.7720633550   -1.6551719031    1.3083306617
C         1.8183477910   -2.5114655261    1.1447226403
H         1.5186925904   -2.3211364057    2.1708497481
H         2.2468312203   -3.5161811189    1.1055887585
N         0.6299237930   -2.4270430596    0.2988497905
C        -0.5068643211   -3.1089500045    0.8918051870
H        -0.5466360484   -2.8630623135    1.9520530538
H        -0.4343519661   -4.1928628665    0.7969670330
C         0.8935798983   -2.9386731454   -1.0412389843
H        -0.0486256857   -3.2782386446   -1.4728301653
H         1.5724219932   -3.7918065977   -1.0079488043
C         1.4233566415   -1.9087675700   -2.0319888411
O         2.2109505240   -2.3122028268   -2.9092736279
O         0.9825839671   -0.7268863108   -1.9453049545
C         2.3829641222    2.0860224616   -0.2594662391
O         3.0634394153    3.0701810270   -0.5997062929
O         1.1184803816    2.0773259083   -0.1668114645
C        -1.8331887285   -2.6464109009    0.2977970549
O        -2.7788672875   -3.4549326696    0.3165831130
O        -1.9058177974   -1.4584890016   -0.1384842446
$end

$rem
 jobtype = opt
 method =  wb97x-d
 basis =  def2-tzvp
 solvent_method = smd
```

```
 mem_static = 1000
 mem_total = 16000
 geom_opt_max_cycles = 600
 geom_opt_tol_gradient = 40
 geom_opt_tol_energy = 20
 geom_opt_tol_displacement = 100
 geom_opt_dmax = 10
 scf_convergence = 8
 max_scf_cycles = 5000
 xc_grid = 2
$end


$molecule
READ
$end

$rem
 jobtype = freq
 method =  wb97x-d
 basis = def2-tzvp
 solvent_method = smd
 mem_static = 1000
 mem_total = 16000
 max_scf_cycles = 5000
 vibman_print =   6
 xc_grid = 2
$end
```

---

mg_opt_freq_m.in

```
$molecule
-2 1
Mg       -0.0124507249    0.0075219432    0.0175835464
O         0.0139254005   -0.0627736841    2.0752511483
C         1.1456925045   -0.0885209254    2.6481248430
O         1.3232549621   -0.2389480008    3.8669426661
C         2.3710360196    0.1409716142    1.7760246012
H         3.2087769549   -0.4323017498    2.1704889313
H         2.6070302463    1.1977272175    1.8937756706
N         2.1599331614   -0.1387509331    0.3606179717
C         2.7062738950    0.9089299218   -0.4829352789
H         2.7451992808    0.5542249731   -1.5101026923
H         3.7099480972    1.2080532360   -0.1839543942
C         2.5651085244   -1.4939057112   -0.0069445450
H         2.6136031816   -1.5505066844   -1.0910673763
```

```
H          3.5532136168   -1.7325356308    0.3877971091
C          1.5412251046   -2.4838715799    0.5129921882
H          1.4596688408   -2.3972576996    1.5930627428
H          1.8507840353   -3.5024900426    0.2773993639
N          0.2364251048   -2.1792720031   -0.0696977603
C         -0.8992725820   -2.6806594357    0.6813613873
H         -0.6872092899   -2.5957451341    1.7444221873
H         -1.1245854830   -3.7217366792    0.4558335727
C          0.1612984420   -2.5478786201   -1.4777860854
H         -0.8581848628   -2.8543061637   -1.7094200962
H          0.8190409978   -3.3859171336   -1.7040159864
C          0.4558440474   -1.4028773583   -2.4324257605
O          0.7884641289   -1.6939692022   -3.5933739564
O          0.2934481977   -0.2234256729   -1.9989232734
C          1.7954279723    2.1333757498   -0.4881091324
O          2.2829479984    3.2138171389   -0.8504413882
O          0.5792725210    1.9631653982   -0.1511059847
C         -2.1378833719   -1.8293766691    0.4192357106
O         -3.2463681463   -2.3386402709    0.6423403495
O         -1.9578120331   -0.6358096355    0.0151833948
$end

$rem
 jobtype = opt
 method =  m05-2x
 basis = def2-tzvp
 solvent_method = smd
 mem_static = 1000
 mem_total = 16000
 geom_opt_max_cycles = 600
 geom_opt_tol_gradient = 40
 geom_opt_tol_energy = 20
 geom_opt_tol_displacement = 100
 geom_opt_dmax = 10
 scf_convergence = 8
 max_scf_cycles = 5000
 xc_grid = 3
 chelpg true
$end


$molecule
READ
$end

$rem
 jobtype = freq
```

```
 method =  m05-2x
 basis = def2-tzvp
 solvent_method = smd
 mem_static = 1000
 mem_total = 16000
 max_scf_cycles = 5000
 vibman_print = 6
 xc_grid = 3
$end
```

---

mg_opt_freq_w.in

```
$molecule
-2 1
Mg       -0.0124507249    0.0075219432    0.0175835464
O         0.0139254005   -0.0627736841    2.0752511483
C         1.1456925045   -0.0885209254    2.6481248430
O         1.3232549621   -0.2389480008    3.8669426661
C         2.3710360196    0.1409716142    1.7760246012
H         3.2087769549   -0.4323017498    2.1704889313
H         2.6070302463    1.1977272175    1.8937756706
N         2.1599331614   -0.1387509331    0.3606179717
C         2.7062738950    0.9089299218   -0.4829352789
H         2.7451992808    0.5542249731   -1.5101026923
H         3.7099480972    1.2080532360   -0.1839543942
C         2.5651085244   -1.4939057112   -0.0069445450
H         2.6136031816   -1.5505066844   -1.0910673763
H         3.5532136168   -1.7325356308    0.3877971091
C         1.5412251046   -2.4838715799    0.5129921882
H         1.4596688408   -2.3972576996    1.5930627428
H         1.8507840353   -3.5024900426    0.2773993639
N         0.2364251048   -2.1792720031   -0.0696977603
C        -0.8992725820   -2.6806594357    0.6813613873
H        -0.6872092899   -2.5957451341    1.7444221873
H        -1.1245854830   -3.7217366792    0.4558335727
C         0.1612984420   -2.5478786201   -1.4777860854
H        -0.8581848628   -2.8543061637   -1.7094200962
H         0.8190409978   -3.3859171336   -1.7040159864
C         0.4558440474   -1.4028773583   -2.4324257605
O         0.7884641289   -1.6939692022   -3.5933739564
O         0.2934481977   -0.2234256729   -1.9989232734
C         1.7954279723    2.1333757498   -0.4881091324
O         2.2829479984    3.2138171389   -0.8504413882
O         0.5792725210    1.9631653982   -0.1511059847
C        -2.1378833719   -1.8293766691    0.4192357106
O        -3.2463681463   -2.3386402709    0.6423403495
```

```
O        -1.9578120331   -0.6358096355    0.0151833948
$end

$rem
 jobtype = opt
 method =  wB97X-D
 basis = def2-tzvp
 solvent_method = smd
 mem_static = 1000
 mem_total = 16000
 geom_opt_max_cycles = 600
 geom_opt_tol_gradient = 40
 geom_opt_tol_energy = 20
 geom_opt_tol_displacement = 100
 geom_opt_dmax = 10
 scf_convergence = 8
 max_scf_cycles = 5000
 xc_grid = 2
 chelpg true
$end


$molecule
READ
$end

$rem
 jobtype = freq
 method =  wB97X-D
 basis = def2-tzvp
 solvent_method = smd
 mem_static = 1000
 mem_total = 16000
 max_scf_cycles = 5000
 vibman_print = 6
 xc_grid = 2
$end

-----------------------------------------------------
 ca_eda_w.in
-----------------------------------------------------
 $molecule
-2 1
--
 2 1
 Ca        0.0568200754   -0.0340507487   -0.0241536722
--
```

```
-4 1
 O        -0.0266333451   -0.0060214852    2.2942122857
 C         1.0269754008    0.0129202493    2.9953916923
 O         1.0573624107   -0.0977604490    4.2344439888
 C         2.3596241929    0.2639608226    2.2856921716
 H         3.1635561425   -0.2111629614    2.8569573893
 H         2.5145038971    1.3433782314    2.3673182739
 N         2.4176607442   -0.1187484848    0.8735315702
 C         3.1674825911    0.8574762557    0.0910602416
 H         3.4115669762    0.4189662405   -0.8797092364
 H         4.1141373734    1.1370271339    0.5660796192
 C         2.9173296471   -1.4862563181    0.6905601916
 H         3.1891485562   -1.6127156332   -0.3580773736
 H         3.8298079155   -1.6539823787    1.2772437380
 C         1.8913877356   -2.5361434014    1.0829434647
 H         1.6046453206   -2.3947792303    2.1256885698
 H         2.3625700878   -3.5253335239    1.0164534456
 N         0.6780693593   -2.4724222800    0.2604912448
 C        -0.4722615912   -3.0916450357    0.9082307927
 H        -0.4477093347   -2.8532230990    1.9745531057
 H        -0.4633957757   -4.1838387038    0.8212005629
 C         0.8878857548   -3.0171302831   -1.0810810247
 H        -0.0606039110   -3.4111916099   -1.4560322841
 H         1.5894218509   -3.8576305915   -1.0630492552
 C         1.3341105049   -2.0166103716   -2.1497508729
 O         1.9332069908   -2.4739861861   -3.1396755917
 O         1.0142557700   -0.8013057071   -1.9967400321
 C         2.3705987271    2.1291552057   -0.2211942199
 O         3.0186263330    3.1575254362   -0.4842999709
 O         1.1058168025    2.0461479272   -0.2403311654
 C        -1.8170416944   -2.5600853818    0.3980975935
 O        -2.8133043262   -3.2867204036    0.5591204716
 O        -1.8429917341   -1.3986316291   -0.1093428981
 $end

 $rem
  jobtype = eda
  method = wb97x-d
  basis = def2-tzvp
  mem_static = 1000
  mem_total = 16000
  solvent_method = smd
 $end

------------------------------------------------------------
 mg_eda_w.in
------------------------------------------------------------
```

```
$molecule
-2 1
--
 2 1
 Mg       -0.0189445819    0.0168856118    0.0029552556
--
-4 1
 O         0.0059357732    0.0539462402    2.0666953554
 C         1.1192677168    0.0565124029    2.6707956290
 O         1.2573714505    0.0395244414    3.9053082702
 C         2.3884653211    0.1278723578    1.8240492753
 H         3.1314435353   -0.5492127077    2.2533000325
 H         2.7767185111    1.1399567901    1.9619064517
 N         2.1959491132   -0.1350066786    0.3964608780
 C         2.7399855870    0.9263867367   -0.4406785973
 H         2.8099384384    0.5688976476   -1.4702186279
 H         3.7421857655    1.2387509233   -0.1326537005
 C         2.5897052603   -1.4944006285    0.0129942865
 H         2.6749552922   -1.5336926682   -1.0739775698
 H         3.5697888367   -1.7626000133    0.4232645051
 C         1.5535070376   -2.5061048344    0.4841153080
 H         1.4573549177   -2.4562572045    1.5697776586
 H         1.8934956444   -3.5174141170    0.2341259076
 N         0.2426874615   -2.2169078213   -0.1065384835
 C        -0.8986497874   -2.7109933035    0.6529204644
 H        -0.6719353043   -2.6464020709    1.7193064090
 H        -1.1376935237   -3.7548077148    0.4292616747
 C         0.1689683890   -2.5815163035   -1.5231226018
 H        -0.8081969643   -3.0209737048   -1.7379240610
 H         0.9132956900   -3.3412529777   -1.7746027711
 C         0.3135749725   -1.4133154654   -2.4968563041
 O         0.5092072377   -1.6863795543   -3.6927709236
 O         0.1881693623   -0.2414664394   -2.0325203989
 C         1.8191836002    2.1528535142   -0.4809782972
 O         2.3167054138    3.2405970779   -0.8091789340
 O         0.5895630873    1.9703752935   -0.2125263301
 C        -2.1472787881   -1.8468625277    0.4343522829
 O        -3.2537365269   -2.3577093572    0.6653267987
 O        -1.9679111967   -0.6429723725    0.0657648319
$end

$rem
 jobtype = eda
 method = wb97x-d
 basis = def2-tzvp
 mem_static = 1000
 mem_total = 16000
```

```
 solvent_method = smd
$end
```

---

cainmg_eda_w.in

```
$molecule
-2 1
--
 2 1
 Ca       -0.0189445819    0.0168856118    0.0029552556
--
-4 1
 O         0.0059357732    0.0539462402    2.0666953554
 C         1.1192677168    0.0565124029    2.6707956290
 O         1.2573714505    0.0395244414    3.9053082702
 C         2.3884653211    0.1278723578    1.8240492753
 H         3.1314435353   -0.5492127077    2.2533000325
 H         2.7767185111    1.1399567901    1.9619064517
 N         2.1959491132   -0.1350066786    0.3964608780
 C         2.7399855870    0.9263867367   -0.4406785973
 H         2.8099384384    0.5688976476   -1.4702186279
 H         3.7421857655    1.2387509233   -0.1326537005
 C         2.5897052603   -1.4944006285    0.0129942865
 H         2.6749552922   -1.5336926682   -1.0739775698
 H         3.5697888367   -1.7626000133    0.4232645051
 C         1.5535070376   -2.5061048344    0.4841153080
 H         1.4573549177   -2.4562572045    1.5697776586
 H         1.8934956444   -3.5174141170    0.2341259076
 N         0.2426874615   -2.2169078213   -0.1065384835
 C        -0.8986497874   -2.7109933035    0.6529204644
 H        -0.6719353043   -2.6464020709    1.7193064090
 H        -1.1376935237   -3.7548077148    0.4292616747
 C         0.1689683890   -2.5815163035   -1.5231226018
 H        -0.8081969643   -3.0209737048   -1.7379240610
 H         0.9132956900   -3.3412529777   -1.7746027711
 C         0.3135749725   -1.4133154654   -2.4968563041
 O         0.5092072377   -1.6863795543   -3.6927709236
 O         0.1881693623   -0.2414664394   -2.0325203989
 C         1.8191836002    2.1528535142   -0.4809782972
 O         2.3167054138    3.2405970779   -0.8091789340
 O         0.5895630873    1.9703752935   -0.2125263301
 C        -2.1472787881   -1.8468625277    0.4343522829
 O        -3.2537365269   -2.3577093572    0.6653267987
 O        -1.9679111967   -0.6429723725    0.0657648319
$end
```

```
$rem
 jobtype = eda
 method = wb97x-d
 basis = def2-tzvp
 mem_static = 1000
 mem_total = 16000
 solvent_method = smd
$end

------------------------------------------------------------
mginca_eda_w.in
------------------------------------------------------------
$molecule
-2 1
--
 2 1
 Mg        0.0568200754   -0.0340507487   -0.0241536722
--
-4 1
 O        -0.0266333451   -0.0060214852    2.2942122857
 C         1.0269754008    0.0129202493    2.2953916923
 O         1.0573624107   -0.0977604490    4.2344439888
 C         2.3596241929    0.2639608226    2.2856921716
 H         3.1635561425   -0.2111629614    2.8569573893
 H         2.5145038971    1.3433782314    2.3673182739
 N         2.4176607442   -0.1187484848    0.8735315702
 C         3.1674825911    0.8574762557    0.0910602416
 H         3.4115669762    0.4189662405   -0.8797092364
 H         4.1141373734    1.1370271339    0.5660796192
 C         2.9173296471   -1.4862563181    0.6905601916
 H         3.1891485562   -1.6127156332   -0.3580773736
 H         3.8298079155   -1.6539823787    1.2772437380
 C         1.8913877356   -2.5361434014    1.0829434647
 H         1.6046453206   -2.3947792303    2.1256885698
 H         2.3625700878   -3.5253335239    1.0164534456
 N         0.6780693593   -2.4724222800    0.2604912448
 C        -0.4722615912   -3.0916450357    0.9082307927
 H        -0.4477093347   -2.8532230990    1.9745531057
 H        -0.4633957757   -4.1838387038    0.8212005629
 C         0.8878857548   -3.0171302831   -1.0810810247
 H        -0.0606039110   -3.4111916099   -1.4560322841
 H         1.5894218509   -3.8576305915   -1.0630492552
 C         1.3341105049   -2.0166103716   -2.1497508729
 O         1.9332069908   -2.4739861861   -3.1396755917
 O         1.0142557700   -0.8013057071   -1.9967400321
 C         2.3705987271    2.1291552057   -0.2211942199
 O         3.0186263330    3.1575254362   -0.4842999709
```

```
O          1.1058168025    2.0461479272   -0.2403311654
C         -1.8170416944   -2.5600853818    0.3980975935
O         -2.8133043262   -3.2867204036    0.5591204716
O         -1.8429917341   -1.3986316291   -0.1093428981
$end

$rem
 jobtype = eda
 method = wb97x-d
 basis = def2-tzvp
 mem_static = 1000
 mem_total = 16000
 solvent_method = smd
$end
```

---

ca_sapt.in

```
molecule {
2 1
Ca         0.0568200754   -0.0340507487   -0.0241536722
--
-4 1
O         -0.0266333451   -0.0060214852    2.2942122857
C          1.0269754008    0.0129202493    2.2953916923
O          1.0573624107   -0.0977604490    4.2344439888
C          2.3596241929    0.2639608226    2.2856921716
H          3.1635561425   -0.2111629614    2.8569573893
H          2.5145038971    1.3433782314    2.3673182739
N          2.4176607442   -0.1187484848    0.8735315702
C          3.1674825911    0.8574762557    0.0910602416
H          3.4115669762    0.4189662405   -0.8797092364
H          4.1141373734    1.1370271339    0.5660796192
C          2.9173296471   -1.4862563181    0.6905601916
H          3.1891485562   -1.6127156332   -0.3580773736
H          3.8298079155   -1.6539823787    1.2772437380
C          1.8913877356   -2.5361434014    1.0829434647
H          1.6046453206   -2.3947792303    2.1256885698
H          2.3625700878   -3.5253335239    1.0164534456
N          0.6780693593   -2.4724222800    0.2604912448
C         -0.4722615912   -3.0916450357    0.9082307927
H         -0.4477093347   -2.8532230990    1.9745531057
H         -0.4633957757   -4.1838387038    0.8212005629
C          0.8878857548   -3.0171302831   -1.0810810247
H         -0.0606039110   -3.4111916099   -1.4560322841
H          1.5894218509   -3.8576305915   -1.0630492552
C          1.3341105049   -2.0166103716   -2.1497508729
```

```
O          1.9332069908   -2.4739861861   -3.1396755917
O          1.0142557700   -0.8013057071   -1.9967400321
C          2.3705987271    2.1291552057   -0.2211942199
O          3.0186263330    3.1575254362   -0.4842999709
O          1.1058168025    2.0461479272   -0.2403311654
C         -1.8170416944   -2.5600853818    0.3980975935
O         -2.8133043262   -3.2867204036    0.5591204716
O         -1.8429917341   -1.3986316291   -0.1093428981

units angstrom
no_reorient
symmetry c1
}

set {
    basis def2-tzvp
}

energy("sapt0-ct")
```

---

mg_sapt.in

```
molecule {
2 1
Mg       -0.0189445819    0.0168856118    0.0029552556
--
-4 1
O          0.0059357732    0.0539462402    2.0666953554
C          1.1192677168    0.0565124029    2.6707956290
O          1.2573714505    0.0395244414    3.9053082702
C          2.3884653211    0.1278723578    1.8240492753
H          3.1314435353   -0.5492127077    2.2533000325
H          2.7767185111    1.1399567901    1.9619064517
N          2.1959491132   -0.1350066786    0.3964608780
C          2.7399855870    0.9263867367   -0.4406785973
H          2.8099384384    0.5688976476   -1.4702186279
H          3.7421857655    1.2387509233   -0.1326537005
C          2.5897052603   -1.4944006285    0.0129942865
H          2.6749552922   -1.5336926682   -1.0739775698
H          3.5697888367   -1.7626000133    0.4232645051
C          1.5535070376   -2.5061048344    0.4841153080
H          1.4573549177   -2.4562572045    1.5697776586
H          1.8934956444   -3.5174141170    0.2341259076
N          0.2426874615   -2.2169078213   -0.1065384835
C         -0.8986497874   -2.7109933035    0.6529204644
H         -0.6719353043   -2.6464020709    1.7193064090
```

```
H        -1.1376935237   -3.7548077148    0.4292616747
C         0.1689683890   -2.5815163035   -1.5231226018
H        -0.8081969643   -3.0209737048   -1.7379240610
H         0.9132956900   -3.3412529777   -1.7746027711
C         0.3135749725   -1.4133154654   -2.4968563041
O         0.5092072377   -1.6863795543   -3.6927709236
O         0.1881693623   -0.2414664394   -2.0325203989
C         1.8191836002    2.1528535142   -0.4809782972
O         2.3167054138    3.2405970779   -0.8091789340
O         0.5895630873    1.9703752935   -0.2125263301
C        -2.1472787881   -1.8468625277    0.4343522829
O        -3.2537365269   -2.3577093572    0.6653267987
O        -1.9679111967   -0.6429723725    0.0657648319

units angstrom
no_reorient
symmetry c1
}

set {
    basis def2-tzvp
}

energy("sapt0-ct")
```

---

cainmg_sapt.in

```
molecule {
2 1
Ca       -0.0189445819    0.0168856118    0.0029552556
--
-4 1
O         0.0059357732    0.0539462402    2.0666953554
C         1.1192677168    0.0565124029    2.6707956290
O         1.2573714505    0.0395244414    3.9053082702
C         2.3884653211    0.1278723578    1.8240492753
H         3.1314435353   -0.5492127077    2.2533000325
H         2.7767185111    1.1399567901    1.9619064517
N         2.1959491132   -0.1350066786    0.3964608780
C         2.7399855870    0.9263867367   -0.4406785973
H         2.8099384384    0.5688976476   -1.4702186279
H         3.7421857655    1.2387509233   -0.1326537005
C         2.5897052603   -1.4944006285    0.0129942865
H         2.6749552922   -1.5336926682   -1.0739775698
H         3.5697888367   -1.7626000133    0.4232645051
C         1.5535070376   -2.5061048344    0.4841153080
```

```
H         1.4573549177   -2.4562572045    1.5697776586
H         1.8934956444   -3.5174141170    0.2341259076
N         0.2426874615   -2.2169078213   -0.1065384835
C        -0.8986497874   -2.7109933035    0.6529204644
H        -0.6719353043   -2.6464020709    1.7193064090
H        -1.1376935237   -3.7548077148    0.4292616747
C         0.1689683890   -2.5815163035   -1.5231226018
H        -0.8081969643   -3.0209737048   -1.7379240610
H         0.9132956900   -3.3412529777   -1.7746027711
C         0.3135749725   -1.4133154654   -2.4968563041
O         0.5092072377   -1.6863795543   -3.6927709236
O         0.1881693623   -0.2414664394   -2.0325203989
C         1.8191836002    2.1528535142   -0.4809782972
O         2.3167054138    3.2405970779   -0.8091789340
O         0.5895630873    1.9703752935   -0.2125263301
C        -2.1472787881   -1.8468625277    0.4343522829
O        -3.2537365269   -2.3577093572    0.6653267987
O        -1.9679111967   -0.6429723725    0.0657648319

units angstrom
no_reorient
symmetry c1
}

set {
    basis def2-tzvp
}

energy("sapt0-ct")
```

---

mginca_sapt.in

```
molecule {
2 1
Mg        0.0568200754   -0.0340507487   -0.0241536722
--
-4 1
O        -0.0266333451   -0.0060214852    2.2942122857
C         1.0269754008    0.0129202493    2.9953916923
O         1.0573624107   -0.0977604490    4.2344439888
C         2.3596241929    0.2639608226    2.2856921716
H         3.1635561425   -0.2111629614    2.8569573893
H         2.5145038971    1.3433782314    2.3673182739
N         2.4176607442   -0.1187484848    0.8735315702
C         3.1674825911    0.8574762557    0.0910602416
H         3.4115669762    0.4189662405   -0.8797092364
```

```
H         4.1141373734    1.1370271339    0.5660796192
C         2.9173296471   -1.4862563181    0.6905601916
H         3.1891485562   -1.6127156332   -0.3580773736
H         3.8298079155   -1.6539823787    1.2772437380
C         1.8913877356   -2.5361434014    1.0829434647
H         1.6046453206   -2.3947792303    2.1256885698
H         2.3625700878   -3.5253335239    1.0164534456
N         0.6780693593   -2.4724222800    0.2604912448
C        -0.4722615912   -3.0916450357    0.9082307927
H        -0.4477093347   -2.8532230990    1.9745531057
H        -0.4633957757   -4.1838387038    0.8212005629
C         0.8878857548   -3.0171302831   -1.0810810247
H        -0.0606039110   -3.4111916099   -1.4560322841
H         1.5894218509   -3.8576305915   -1.0630492552
C         1.3341105049   -2.0166103716   -2.1497508729
O         1.9332069908   -2.4739861861   -3.1396755917
O         1.0142557700   -0.8013057071   -1.9967400321
C         2.3705987271    2.1291552057   -0.2211942199
O         3.0186263330    3.1575254362   -0.4842999709
O         1.1058168025    2.0461479272   -0.2403311654
C        -1.8170416944   -2.5600853818    0.3980975935
O        -2.8133043262   -3.2867204036    0.5591204716
O        -1.8429917341   -1.3986316291   -0.1093428981

units angstrom
no_reorient
symmetry c1
}

set {
    basis def2-tzvp
}

energy("sapt0-ct")
```

---

methane.in

```
$rem
   JOBTYPE Opt
   EXCHANGE B3LYP
   BASIS 6-31G(d)
   GUI=2
$end

$comment
Title
```

```
$end

$molecule
   0 1
   C       -2.42702        0.46905        0.00000
   H       -1.35702        0.46905        0.00000
   H       -2.78368        0.00929        0.89795
   H       -2.78369       -0.07872       -0.84714
   H       -2.78369        1.47657       -0.05081
$end
-----------------------------------------------------
```

A.3 Processing Trials of Data in Chapter 5

The current data processing steps are shown in figures A.15–A.17. The first two figures show how we identify the solvent peaks and the HPTS ground state bands. Then the negative transient at 100 ps, before the UV pump, shows where those solvent bands lie and where to expect the ground state bleaches. Spectrum at 0 ns then reveals the new HPTS* peaks besides the bleaches and solvent bands. Two of the figures (A.18–A.19) are Gaussian fitting spectra after baselining. Those figures show that the two Gaussian model reliably captures the peak shape of the HCOOH band and we can extract the kinetic trend from them.

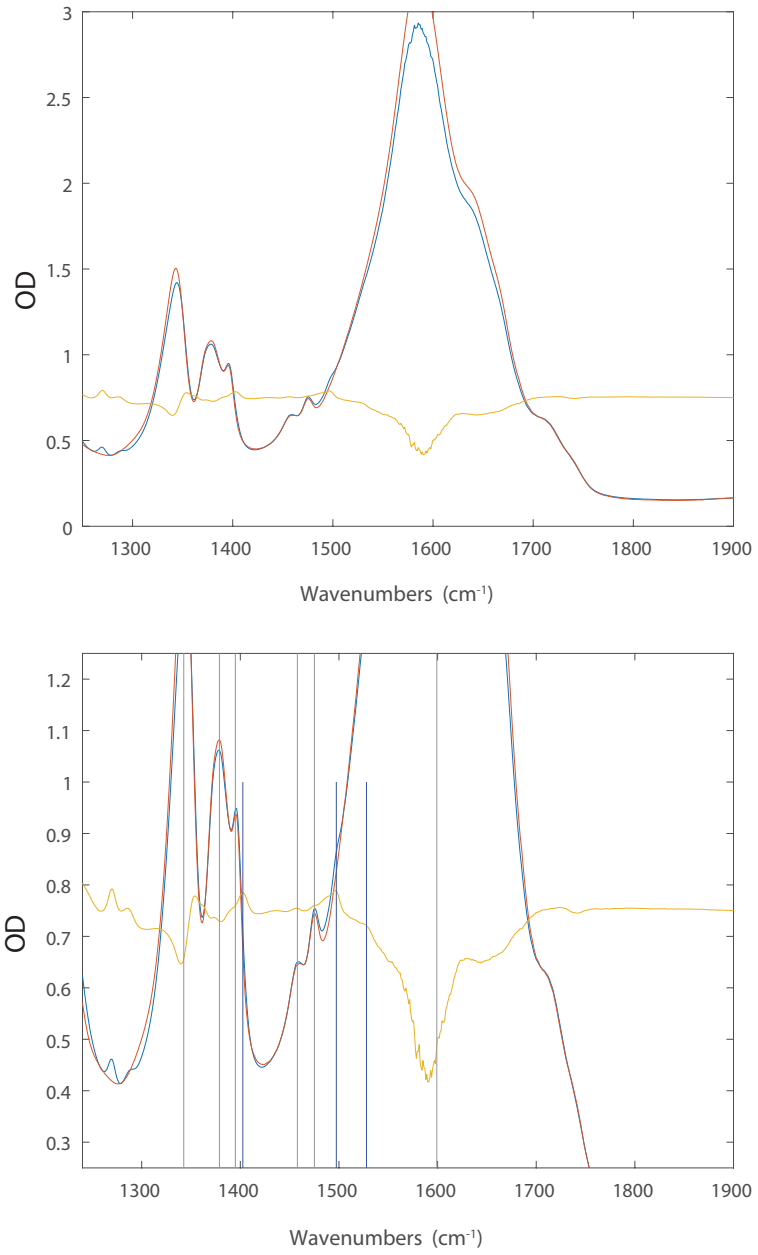

Figure A.15: Subtracting the EAF spectrum (red) from the EAF–HPTS spectrum (blue) gives the solvent shapes and the likely HPTS peaks in the spectrum. The bottom figure marks the solvent band frequencies (gray) and the HPTS ground state peaks (blue).

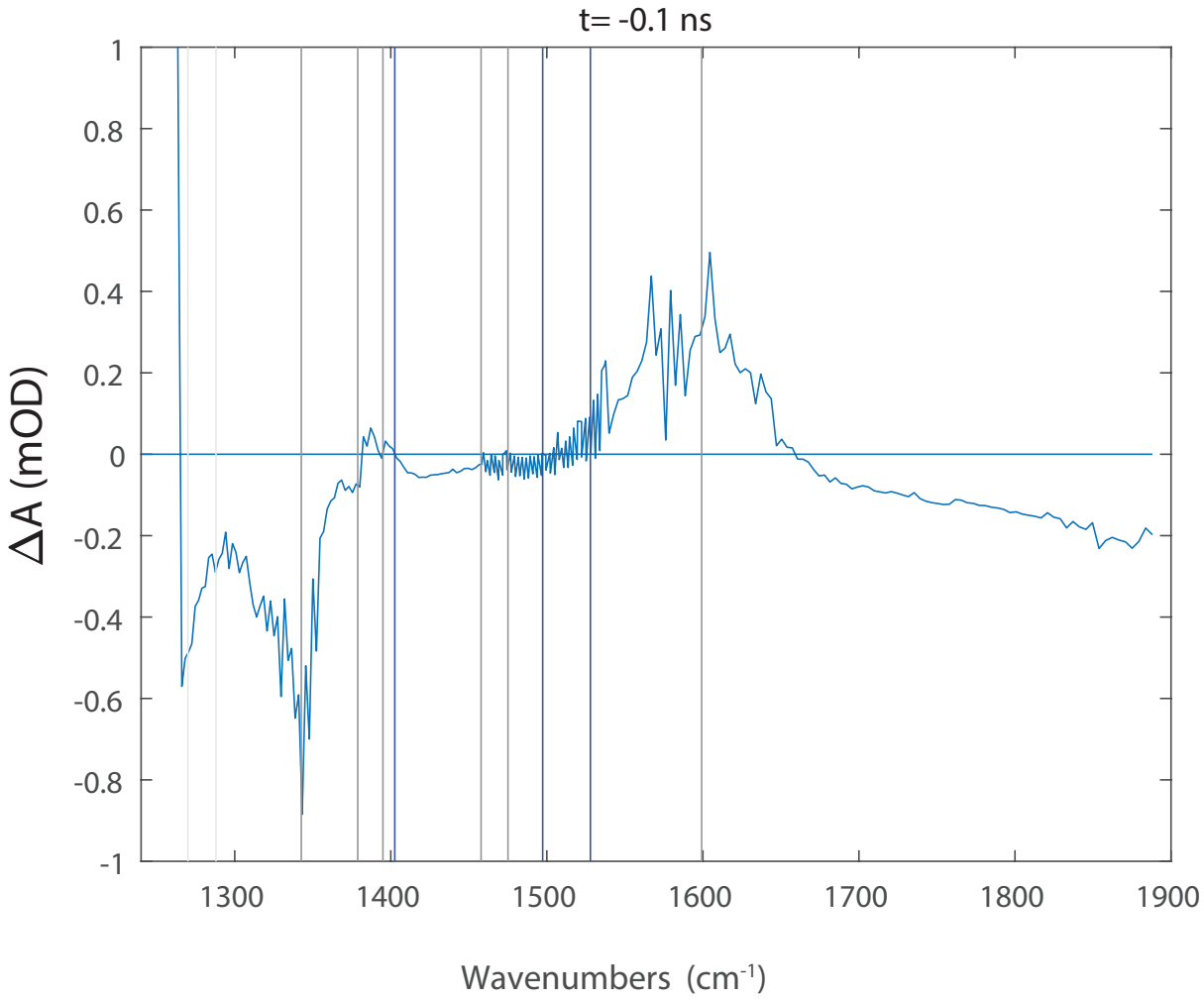

Figure A.16: Spectra at 100 ps before pumping shows the solvent bands (gray lines). The expected HPTS bleaches after UV pump are indicated by blue lines.

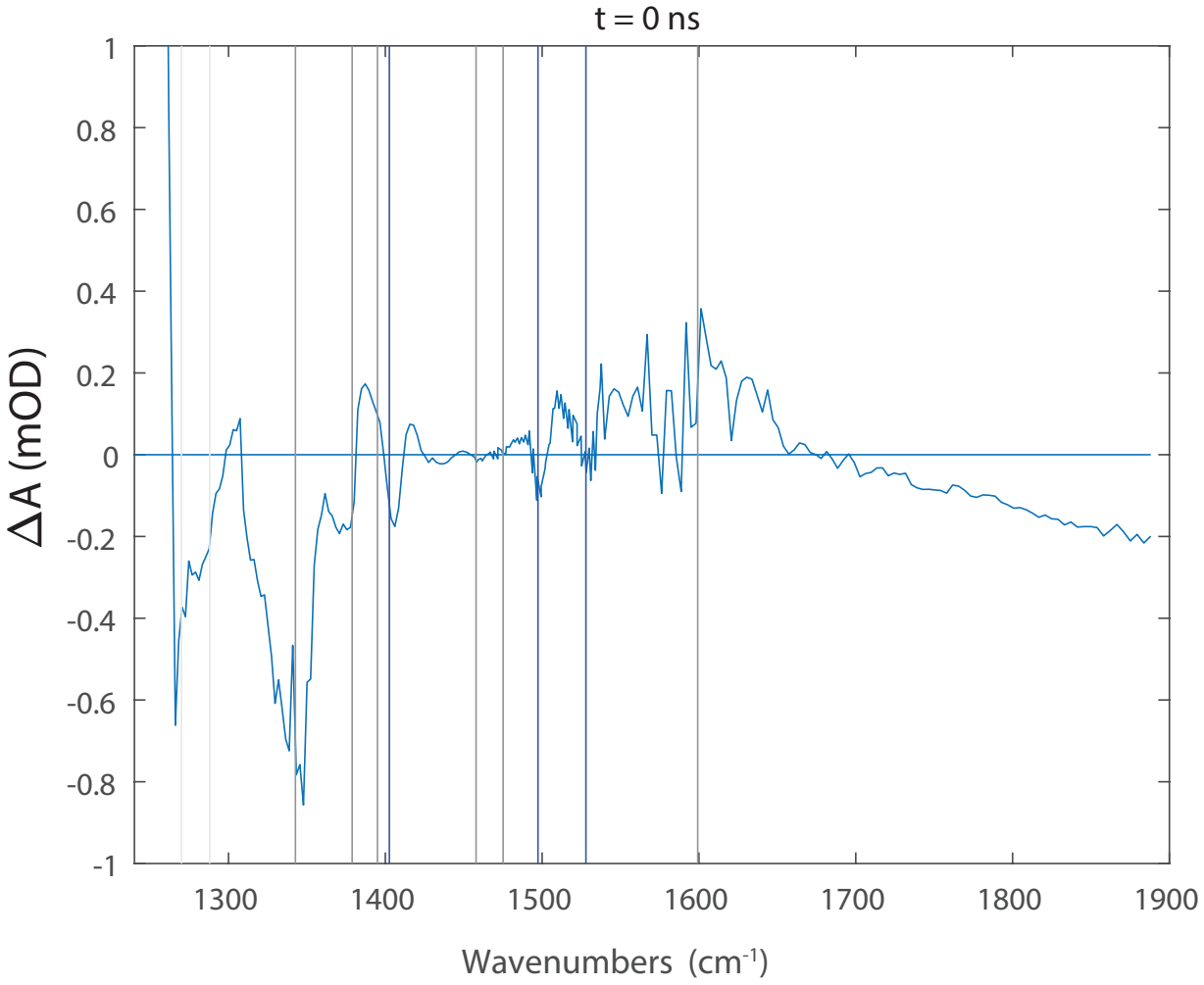

Figure A.17: Markers indicate the solvent and HPTS ground state bands in the time zero transient spectrum. Spectrum identifies the HPTS* bands at time zero as the new excited state absorptions.

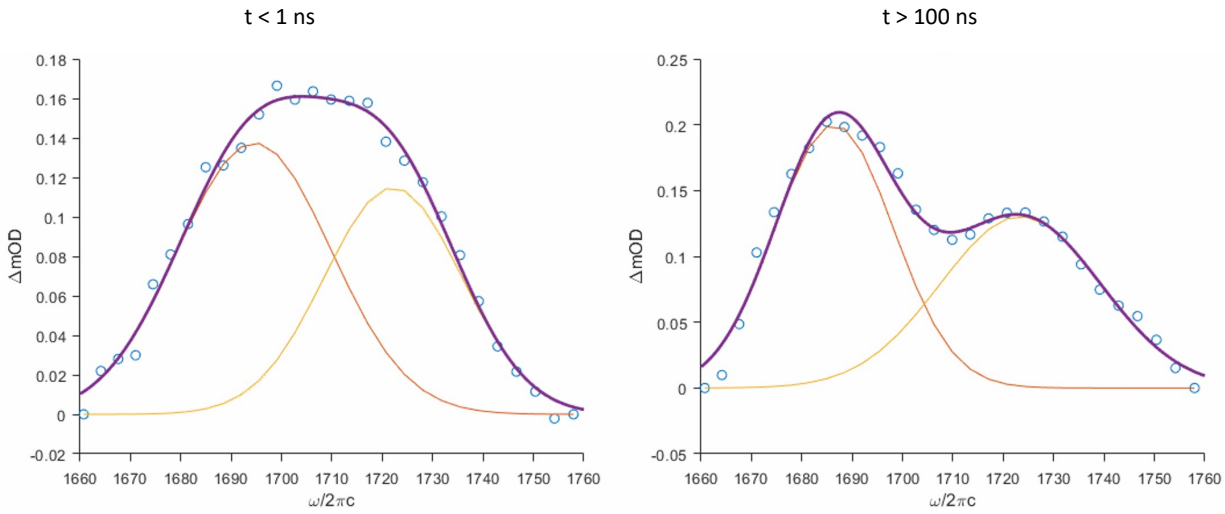

Figure A.18: The two peaks of HCOOH are fitted to two Gaussians. Here blue dots are raw data and underneath lie the Gaussian peaks. At early times the broad peakshape (purple) is captured, while at later times the narrowing of the bands are observed.

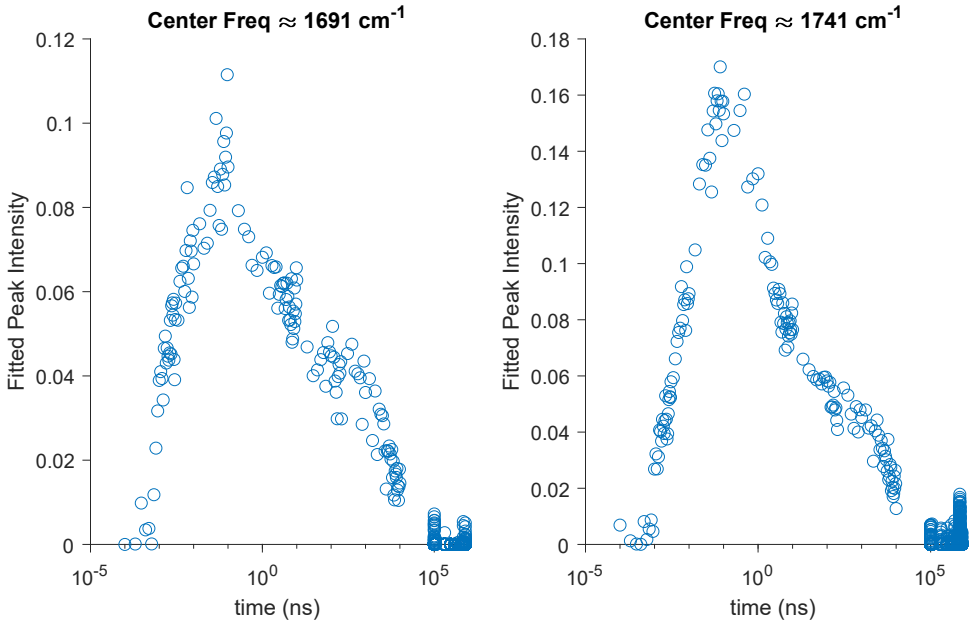

Figure A.19: A double Gaussian was used to model the two peaks HCOOH band. For the peak fitting with KPP, a sum of the amplitudes of these two peaks were used.

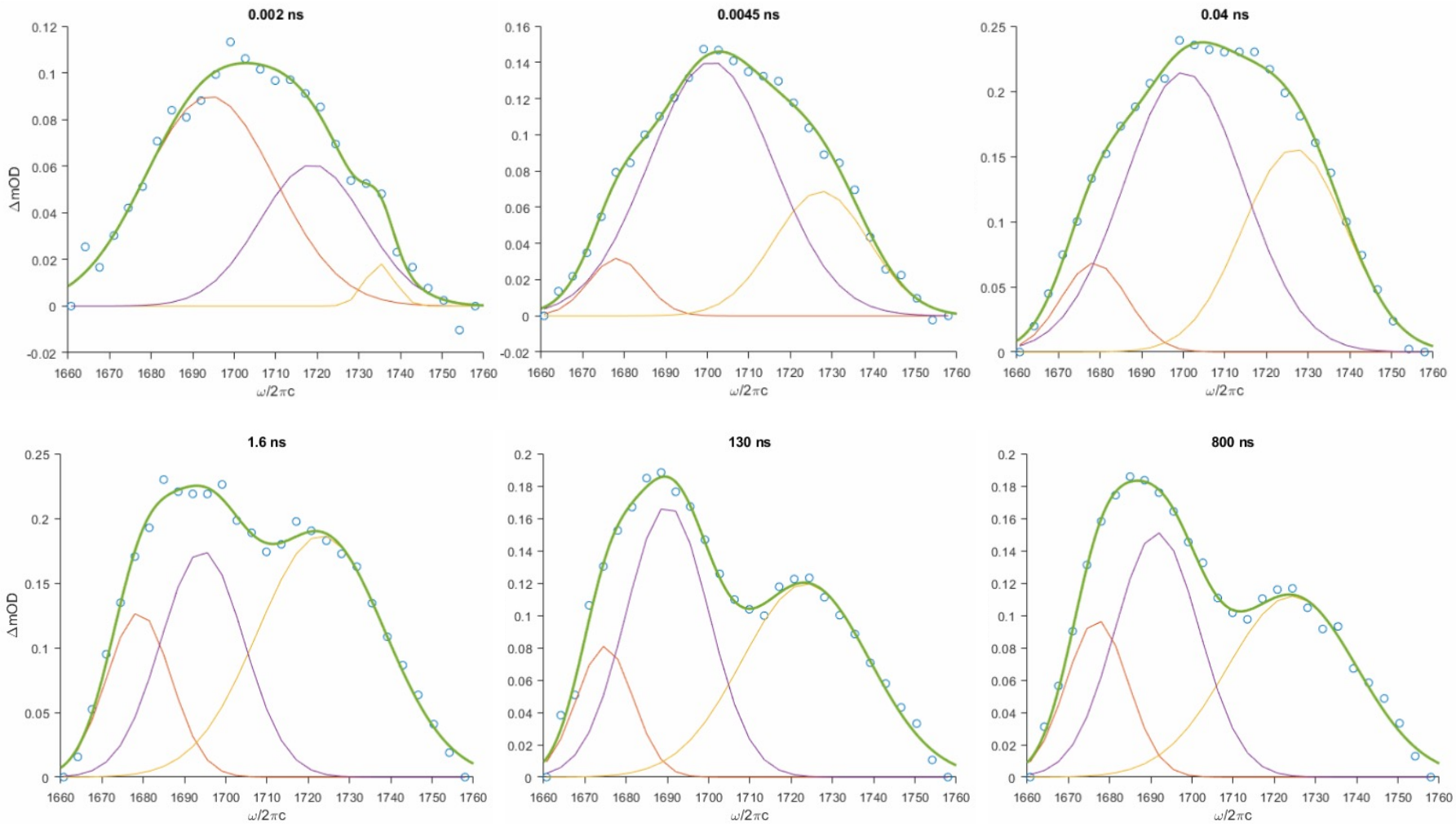

Figure A.20: A triple Gaussian fit on the HCOOH band. The top row represents the early time points before 1 ns, while the bottom row are long time points >1 ns. The top row Gaussians show that the HCOOH peak grows on top of the broad wings of the solvent band before a nanosecond. At long time points the solvent shape over evaluates the HCOOH band amplitude.

Table A.7: Molar extinction coefficients ($M^{-1}cm^{-1}$) for four species.

1404 cm^{-1}	1389 cm^{-1}	1505 cm^{-1}	1720 cm^{-1}
HPTS	HPTS*	PTS^-	HCOOH
165.4 (aq)	9.77	875.9 (aq)	342.0 (aq)

A rough estimate of the extinction coefficients of four species is presented (Table A.7). Three of the approximations are in aqueous solution and the HPTS* species is then evaluated from the measured EAF data. With concentration dependent FTIR experiments of HPTS, PTS^-, and HCOOH aqueous solutions, respectively, their individual molar extinction coefficients were estimated. The HPTS* species concentration was determined using the UV-Vis absorption of HPTS at 400 nm.[276] Using that as the estimate of concentration of the HPTS* 1390 cm^{-1} band (∼18 mM), we then used the highest point of a Gaussian fitted absorbance to obtain the absorbance for HPTS* and calculated the HPTS* extinction coefficient of the 1390 cm^{-1} band. We were unable to calculate the PTS^{*-} extinction coefficient with the current data as a knowledge of concentration in the excited state is necessary, that is unavailable as of yet. Future experiments will try to ascertain the extinction coefficients of all the species in EAF, so that we can observe and model the concentration changes with the ongoing kinetics in each species accurately.

Different pre-processing trials that were initially applied to extract the time evolution of each peak are given here (Figure A.21–A.25). In all the prior attempts to extract the observed kinetics, we first 'pre-processed' the time–resolved data, which includes baseline correction followed by normalization of the maximum peak intensity for a chosen band. After 'pre-processing', the data was fit with kinetic equations for our EAF–HPTS proton reaction cycle generated using KPP. We applied a Gaussian band fitting as an extraction method as well as peak integration on the individual bands to extract the kinetic information from the regions of interest. In two more processing trials, these two methods were applied after

baselining the respective peaks. Another trial with simultaneous Gaussian peak fitting of the spectral region 1370–1540 cm^{-1} was also tried but failed to yield any result. Initially, we also applied SVD to extract the spectral and kinetic components (data not shown). The SVD trials did not produce anything useful as the solvent band correction was not accounted for properly.

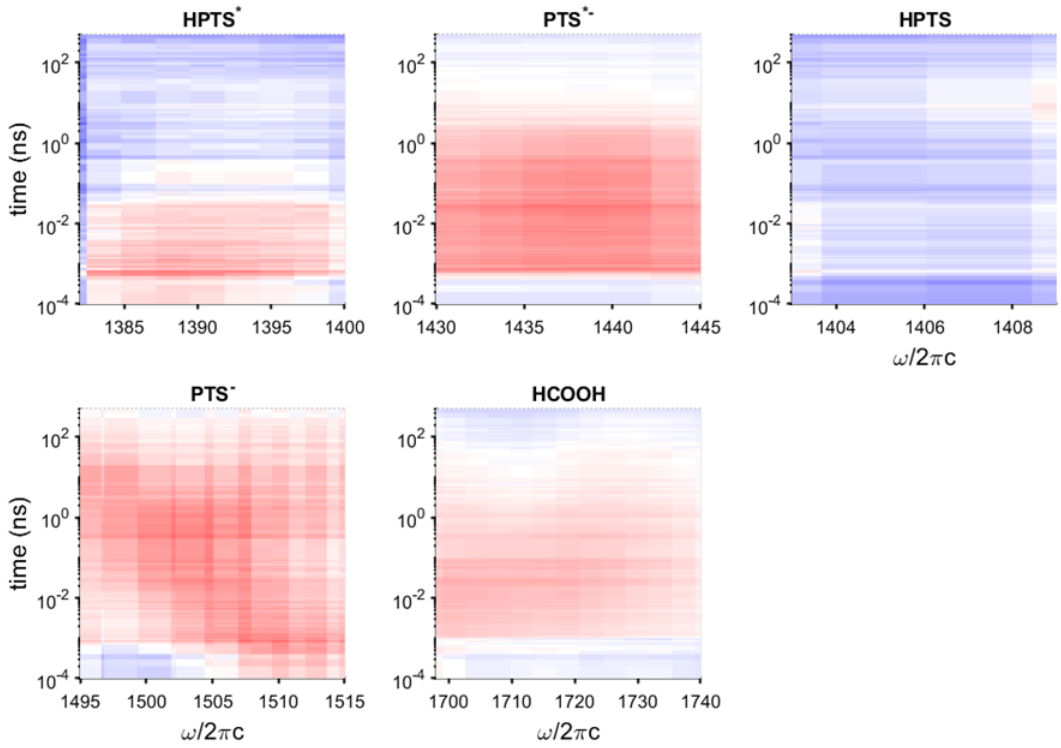

Figure A.21: The ranges for the five peaks that were used in intial baselining and peak integration trials. The HCOOH figure represents the baselined range in the current processing.

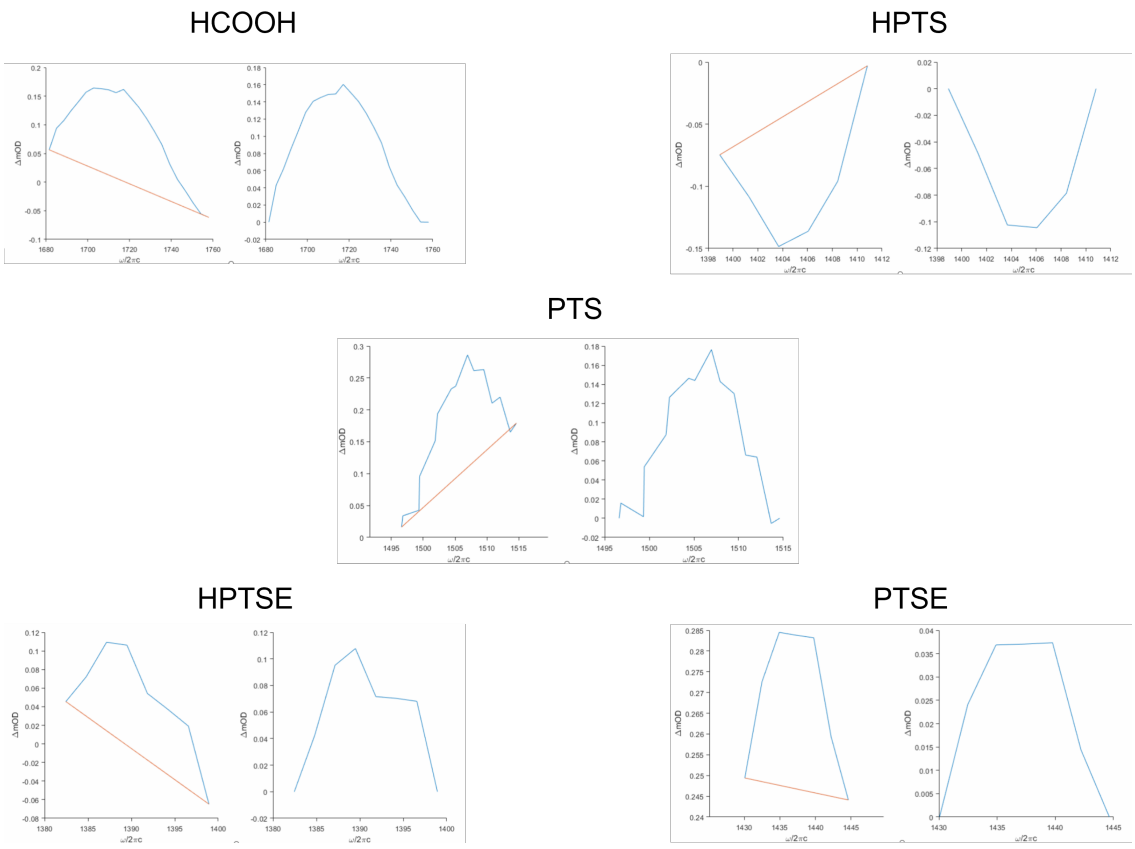

Figure A.22: Baseline for the chosen peaks at different timepoints. In the current processing, we baseline the HCOOH peak before fitting it with two Gaussians. All the other baselines in this figure are example of prior effort to extract peak kinetics.

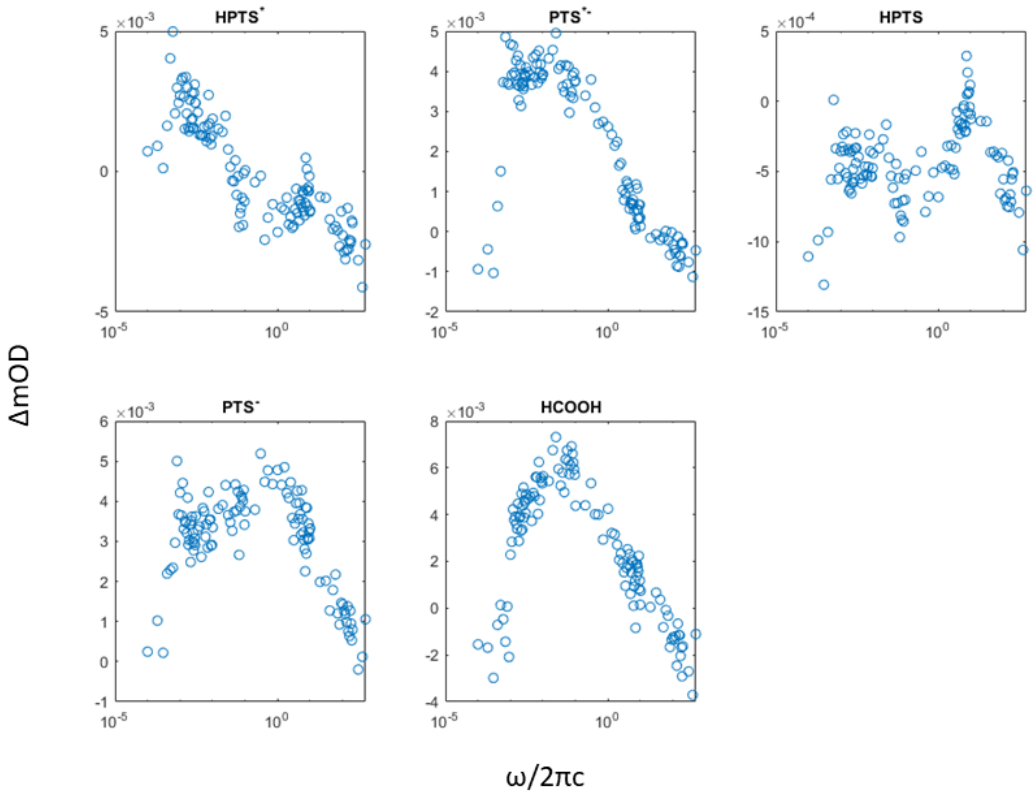

Figure A.23: Integrated area under each peak range showing time evolution of each band. This method shows a very different kinetic trend than the current method.

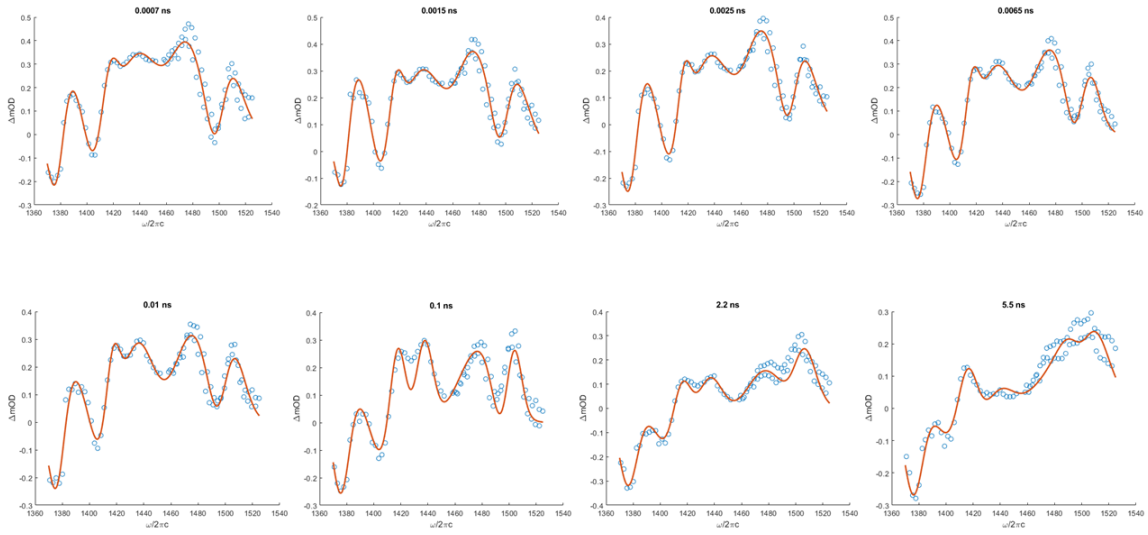

Figure A.24: Applying Gaussian series to fit the 1370-1550cm^{-1}frequency range.

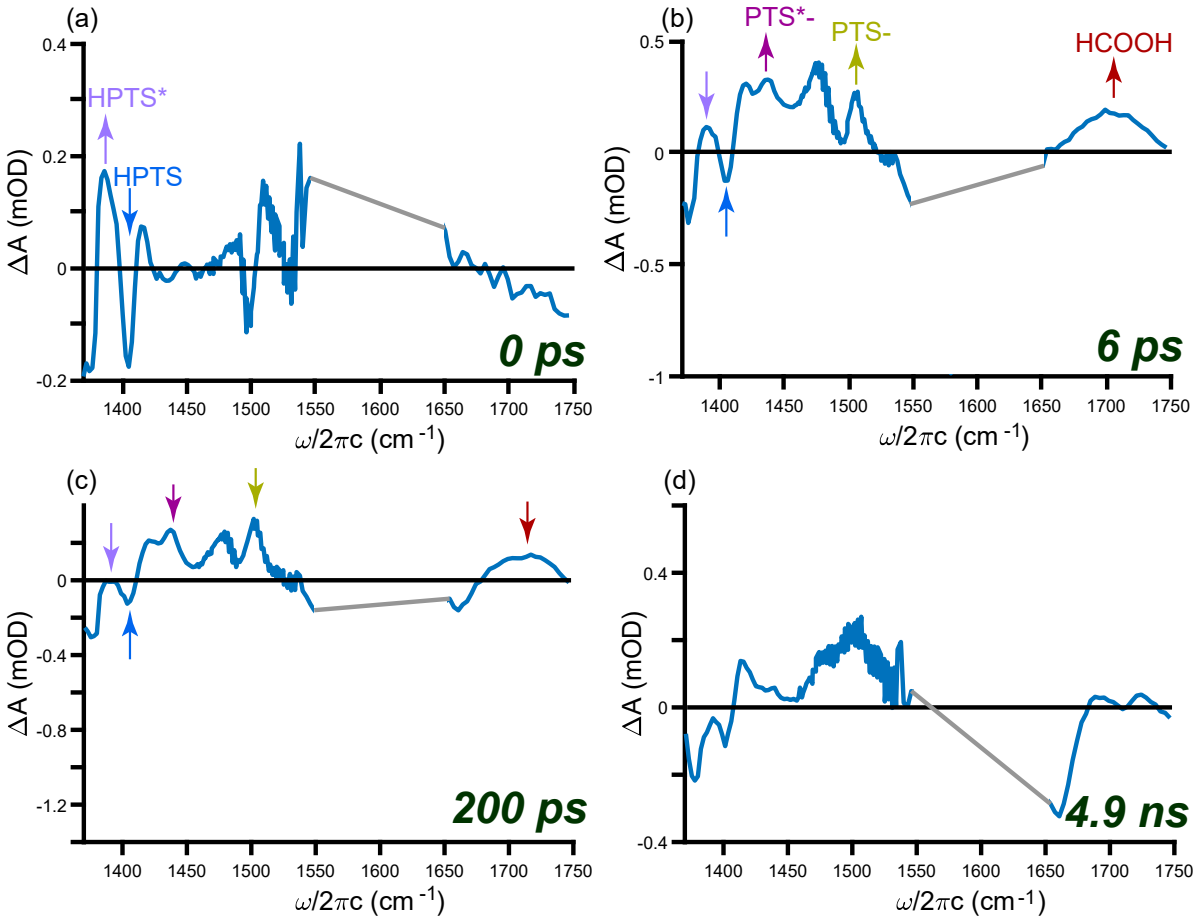

Figure A.25: In the old assignment scheme, only some of the peaks are labeled in the EAF–HPTS system: HPTS ground state bleach (blue), HPTS*transient absorption (violet), PTS^-transient absorption (purple), PTS^{*-}transient absorption (green), and the transient HCOOH peak (red). All the transient peaks rise and decay with time. Formic acid lasts in the EAF system longer than ∼200 ps, disappearing by ∼5 ns. The region 1550–1650 cm^{-1} is the solvent region with no probe light falling in the detector.

A.4 Assignment Instruction Worksheet

This section contains the reference sheet and worksheet template used for every assignment in the curriculum described in chapter 6.

Instructor Sheet of Reference:

1. Pick the secondary article of choice for an assignment.
2. Pick 4 primary research articles related to the secondary article of choice.
3. Rank order the 4 primary research papers according to relatedness to the associated secondary article.
4. Name the 4 articles or their 'doi's in the 4 top right corner boxes of the form.
5. Pick a 'doi' from rank-2 to rank-4 and put them in Q3 and Q4 reference boxes as per your discretion.
6. Place the paper with rank-1 in the reference box for Q5.

Literature Assignment Worksheet:

Name:
Email:

1. List three main points that the secondary article is aiming to convey to readers.

2. List the furnished primary literature articles in order (highest to lowest) of relevance to the secondary article. List using the first author's last name.

3. Answer the following questions based on the article:
   a. Copy and paste the sentence(s) from the paper describing the hypothesis(es).

   b. List the two main goals (in your own words) of the author's work described in the article that relate to the secondary article.

   c. Explain your ranking of this paper in question #2 in terms of relatedness to the assigned secondary article?

4. Answer the following questions based on the article:
   a. Copy and paste the sentence(s) from the paper describing the hypothesis(es).

1

Figure A.26: The first page of the worksheet assignment.

b. List the two main goals (in your own words) of the author's work described in the article that relate to the secondary article.

c. Explain your ranking of this paper in question #2 in terms of relatedness to the assigned secondary article?

5. Answer the following questions for the paper:

a. Copy and paste the sentence(s) from the paper describing the hypothesis(es).

b. List the two main goals of the author's work described in the paper that relate to the secondary article

c. Rank order the category of this article in terms of: "Making", "Modeling", or "Measuring"? Explain the ranking in your own words.

Category	Making	Modeling	Measuring
Order of precedence			
Explanation			

d. Copy and paste the specific sentence(s) from the paper that provide supportive evidence for your answer to question 5c).

2

Figure A.27: The second page of the worksheet assignment.

e. List the three new terms you learned from this work. Write down the definitions in the next column. (Reference your source of definition. List the references in the row below each new term.)

New term	Definition:
Ref 1:	
Ref 2:	
Ref 3:	

f. List two conclusions of this work made by the authors (must be in your own words)? Did the conclusions support or reject the hypothesis?

Conclusions	Support/Rejects

g. Copy and paste the specific sentence(s) from the paper that provide supportive evidence for your answer to question 5f).

h. List one future direction suggested by the author (must be in your own words)?

i. Copy and paste the specific sentence from the paper that provides supportive evidence for your answer to question 5h).

3

Figure A.28: The third page of the worksheet assignment.

Bibliography

[1] Bosmans, F.; Martin-Eauclaire, M.-F.; Swartz, K. J. Deconstructing voltage sensor function and pharmacology in sodium channels. *Nature* **2008**, *456*, 202–208.

[2] Dudev, T.; Lim, C. Factors Governing the Na(+) vs K(+) Selectivity in Sodium Ion Channels. *J. Am. Chem. Soc.* **2010**, *132*, 2321–2332.

[3] Marrone, T. J.; Merz Jr., K. M. Molecular Recongnition of Potassium Ion by the Naturally Occuring Antibiotic Ionophore Nonactin. *J. Am. Chem. Soc.* **1992**, *113*, 7542–7549.

[4] Doyle, D. A.; Cabral, J. M.; Pfuetzner, R. A.; Kuo, A.; Gulbis, J. M.; Cohen, S. L.; Chait, B. T.; MacKinnon, R. The Structure of the Potassium Channel: Molecular Basis of K+ Conduction and Selectivity. *Science* **1998**, *280*, 69–77.

[5] Domene, C.; Klein, M. L.; Branduardi, D.; Gervasio, F. L.; Parrinello, M. Conformational Changes and Gating at the Selectivity Filter of Potassium Channels. *J. Am. Chem. Soc.* **2008**, *130*, 9474–9480.

[6] Thompson, A. N.; Posson, D. J.; Parsa, P. V.; Nimigean, C. M. Molecular Mechanism of pH Sensing in KcsA Potassium Channels. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 6900–6905.

[7] Treptow, W.; Tarek, M. K+ Conduction in the Selectivity Filter of Potassium Channels Is Monitored by the Charge Distribution along Their Sequence. *Biophys. J.* **2006**, *106*, L81–L83.

[8] Moroz, O. V.; Antson, A. A. Structure of the Human S100A12 ± Copper Complex : Implications for Host-Parasite Defence. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **2003**, 859–867.

[9] Soragni, A.; Zambelli, B.; Mukrasch, M. D.; Biernat, J.; Jeganathan, S.; Griesinger, C.; Ciurli, S.; Mandelkow, E.; Zweckstetter, M. Structural Characterization of Binding of Cu(II) to Tau Protein. *Biochemistry* **2008**, *47*, 10841–10851.

[10] Bush, A. I. Metals and neuroscience. *Curr. Opin. Chem. Biol.* **2000**, *4*, 184–191.

[11] Jomova, K.; Valko, M. Advances in Metal-Induced Oxidative Stress and Human Disease. *Toxicology* **2011**, *283*, 65–87.

[12] Valko, M.; Morris, H.; Cronin, M. T. D. Metals, Toxicity and Oxidative Stress. *Curr. Med. Chem.* **2005**, *12*, 1161–1208.

[13] Raymond, K. N.; Dertz, E. A.; Kim, S. S. Enterobactin: An Archetype for Microbial Iron Transport. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3584–3588.

[14] Schafer, B. W.; Fritschy, J.-M.; Murmann, P.; Troxler, H.; Durussel, I.; Heizmann, C. W.; Cox, J. A. Brain S100A5 is a Novel Calcium-, Zinc-, and Copper Ion-Binding Protein of the EF-hand Superfamily. *J. Biol. Chem.* **2000**, *275*, 30623–30630.

[15] Ostendorp, T.; Diez, J.; Heizmann, C. W.; Fritz, G. The Crystal Structures of Human S100B in the Zinc- and Calcium-Loaded State at Three pH Values Reveal Zinc Ligand Swapping. *Biochim. Biophys. Acta - Mol. Cell Res.* **2011**, *1813*, 1083–1091.

[16] Liuzzi, J. P.; Cousins, R. J. Mammalian Zinc Transporters. *Annu. Rev. Nutr.* **2004**, *24*, 151–172.

[17] Capozzi, F.; Casadei, F.; Luchinat, C. EF-hand Protein Dynamics and Evolution of Calcium Signal Transduction: An NMR View. *J. Biol. Inorg. Chem.* **2006**, *11*, 949–962.

[18] Johnson, C. K. Calmodulin, Conformational States, and Calcium Signaling. A Single-Molecule Perspective. *Biochemistry* **2006**, *45*, 14233–14246.

[19] Cates, M. S.; Teodoro, M. J.; Phillips Jr., G. N. Molecular Mechanisms of Calcium and Magnesium Binding to Parvalbumin. *Biophys. J.* **2002**, *82*, 1133–1146.

[20] Reid, R. E.; Hodges, R. S. Co-operativity and Calcium-Magnesium Binding to Troponin-C and Muscle Calcium-Binding Parvalbumin - An Hypothesis. *J. Theor. Biol.* **1980**, *84*, 401–444.

[21] Tottey, S.; Harvie, D. R.; Robinson, N. J. Understanding How Cells Allocate Metals Using Metal Sensors and Metallochaperones. *Acc. Chem. Res.* **2005**, *38*, 775–783.

[22] Zoroddu, M. A.; Aaseth, J.; Crisponi, G.; Medici, S.; Peana, M.; Nurchi, V. M. The Essential Metals for Humans: A Brief Overview. *J. Inorg. Biochem.* **2019**, *195*, 120–129.

[23] Lucia, B.; Ivano, B. *Met. Ions Life Sci.*; Vol. 12; pp 451–478.

[24] Jackson, G. S.; Murray, I.; Hosszu, L. L. P.; Gibbs, N.; Waltho, J. P.; Clarke, A. R.; Collinge, J. Location and Properties of Metal-Binding Sites on the Human Prion Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 8531–8535.

[25] Stryer, L.; Berg, J. M.; Tymoczko, J. L.; Gatto, G. J. *Biochemistry*; W. H. Freeman, 2019.

[26] Voet, D.; Voet, J. G. *Biochemistry*; Wiley, 2004.

[27] Alberts, B. *Molecular Biology of the Cell*; CRC Press, 2017.

[28] Nelson, D. L.; Lehninger, A. L.; Cox, M. M. *Lehninger Principles of Biochemistry*; W. H. Freeman, 2005.

[29] Hamm, P.; Zanni, M. T. *Concepts and Methods of 2D Infrared Spectroscopy*; Cambridge University Press: New York, NY, 2011.

[30] Yoshizawa, M.; Xu, W.; Angell, C. A. Ionic Liquids by Proton Transfer: Vapor Pressure, Conductivity, and the Relevance of ΔpK_a from Aqueous Solutions. *J. Am. Chem. Soc.* **2003**, *125*, 15411–9.

[31] Miran, M. S.; Kinoshita, H.; Yasuda, T.; Susan, M. A. B. H.; Watanabe, M. Physicochemical Properties Determined by ΔpK_a for Protic Ionic Liquids Based on an Organic Super-Strong Base with Various Brønsted Acids. *Phys. Chem. Chem. Phys.* **2012**, *14*, 5178.

[32] Angell, C. A.; Ansari, Y.; Zhao, Z. Ionic Liquids : Past , present and future. **2012**, 9–27.

[33] Belieres, J.-P.; Angell, C. A. Protic Ionic Liquids: Preparation, Characterization, and Proton Free Energy Level Representation †. *J. Phys. Chem. B* **2007**, *111*, 4926–4937.

[34] Nakamoto, H.; Watanabe, M. Brønsted acid-base ionic liquids for fuel cell electrolytes. *Chem. Commun. (Camb).* **2007**, 2539–41.

[35] Smith, D. E.; Walsh, D. A. The Nature of Proton Shuttling in Protic Ionic Liquid Fuel Cells. *Adv. Energy Mater.* **2019**, *9*, 1900744.

[36] Kim, Y. S.; Hochstrasser, R. M. Applications of 2D IR Spectroscopy to Peptides, Proteins, and Hydrogen-Bond Dynamics. *J. Phys. Chem. B* **2009**, *113*, 8231–8251.

[37] Woutersen, S.; Hamm, P. Nonlinear Two-Dimensional Vibrational Spectroscopy of Peptides. *J. Phys. Condens. Matter* **2002**, *14*, R1035–R1062.

[38] Thielges, M. C.; Chung, J. K.; Fayer, M. D. Protein Dynamics in Cytochrome P450 Molecular Recognition and Substrate Specificity Using 2D IR Vibrational Echo Spectroscopy. *J. Am. Chem. Soc.* **2011**, *133*, 3995–4004.

[39] Le Sueur, A. L.; Horness, R. E.; Thielges, M. C. Applications of two-dimensional infrared spectroscopy. *Analyst* **2015**, *140*, 4336–4349.

[40] Maekawa, H.; Toniolo, C.; Broxterman, Q. B.; Ge, N.-h. Two-Dimensional Infrared Spectral Signatures of 3_{10} - and α-Helical Peptides. **2007**, 3222–3235.

[41] Cheatum, C. M.; Tokmakoff, A.; Knoester, J. Signatures of Beta-Sheet Secondary Structures in Linear and Two-Dimensional Infrared Spectroscopy. *J. Chem. Phys.* **2004**, *120*, 8201–8215.

[42] DeFlores, L. P.; Ganim, Z.; Nicodemus, R. A.; Tokmakoff, A. Amide I'-II' 2D IR Spectroscopy Provides Enhanced Protein Secondary Structural Sensitivity. *J. Am. Chem. Soc.* **2009**, *131*, 3385–3391.

[43] Ganim, Z.; Chung, H. S.; Smith, A. W.; DeFlores, L. P.; Jones, K. C.; Tokmakoff, A. Amide I Two-Dimensional Infrared Spectroscopy of Proteins. *Acc. Chem. Res.* **2008**, *41*, 432–441.

[44] Demirdöven, N.; Cheatum, C. M.; Chung, H. S.; Khalil, M.; Knoester, J.; Tokmakoff, A. Two-Dimensional Infrared Spectroscopy of Antiparallel Beta-Sheet Secondary Structure. *J. Am. Chem. Soc.* **2004**, *126*, 7981–7990.

[45] Ghosh, A.; Tucker, M. J.; Gai, F. 2D IR Spectroscopy of Histidine: Probing Side-chain Structure and Dynamics via Backbone Amide Vibrations. *J. Phys. Chem. B* **2014**, *118*, 7799–805.

[46] Minnes, L.; Shaw, D. J.; Cossins, B. P.; Donaldson, P. M.; Greetham, G. M.; Towrie, M.; Parker, A. W.; Baker, M. J.; Henry, A. J.; Taylor, R. J.; Hunt, N. T. Quantifying Secondary Structure Changes in Calmodulin Using 2D-IR Spectroscopy. *Anal. Chem.* **2017**, *89*, 10898–10906.

[47] Ghosh, A.; Ostrander, J. S.; Zanni, M. T. Watching Proteins Wiggle: Mapping Structures with Two-Dimensional Infrared Spectroscopy. *Chem. Rev.* **2017**, *117*, 10726–10759.

[48] Grabarek, Z. Structure of a Trapped Intermediate of Calmodulin: Calcium Regulation of EF-hand Proteins from a New Perspective. *J. Mol. Biol.* **2005**, *346*, 1351–1366.

[49] Grabarek, Z. Structural Basis for Diversity of the EF-hand Calcium-binding Proteins. *J. Mol. Biol.* **2006**, *359*, 509–525.

[50] Grabarek, Z. Insights into Modulation of Calcium Signaling by Magnesium in Calmodulin, Troponin C and Related EF-Hand Proteins. *Biochim. Biophys. Acta - Mol. Cell Res.* **2011**, *1813*, 913–921.

[51] Johnson, C. N.; Damo, S. M.; Chazin, W. J. EF-Hand Calcium-Binding Proteins. *eLS* **2014**,

[52] Suzuki, N.; Imai, L. F.; Kato, Y.; Nagata, K.; Ohashi, Y.; Kuchitsu, K.; Tanokura, M.; Sakamoto, A.; Nara, M.; Nakano, M.; Yonezawa, N. Coordination Structures of Mg 2+ and Ca 2+ in Three Types of Tobacco Calmodulins in Solution: Fourier-Transform Infrared Spectroscopic Studies of Side-Chain COO - Groups. *Biopolymers* **2013**, *99*, 472–483.

[53] DePalma, J. W.; Kelleher, P. J.; Tavares, L. C.; Johnson, M. A. Coordination-Dependent Spectroscopic Signatures of Divalent Metal Ion Binding to Carboxylate Head Groups: H2- and He-Tagged Vibrational Spectra of $M2+ \cdot RCO2^{-}$ (M = Mg and Ca, R = -CD3, -CD2CD3) Complexes. *J. Phys. Chem. Lett.* **2017**, *8*, 484–488.

[54] Nara, M.; Tasumi, M.; Tanokura, M.; Hiraoki, T.; Yazawa, M.; Tsutsumi, A. Infrared studies of interaction between metal ions and Ca 2+ -binding proteins. Marker bands

for identifying the types of coordination of the side-chain COO- groups to metal ions in pike parvalbumin (pI = 4.10). *FEBS Lett.* **1994**, *349*, 84–88.

[55] Gautier, A.; Kirkpatrick, J. P.; Nietlispach, D. Solution-state NMR spectroscopy of a seven-helix transmembrane protein receptor: Backbone assignment, secondary structure, and dynamics. *Angew. Chem.-Int. Ed.* **2008**, *47*, 7297–7300.

[56] Teleman, A.; Drakenberg, T.; Forsen, S. Kinetics of ca2+ binding to calmodulin and its tryptic fragments studied by ca-43-nmr. *Biochim. Biophys. Acta.* **1986**, *873*, 204–213.

[57] Han, S.; Ba, Y. Determination of the concentrations of metal cations in aqueous solutions using proton NMR spectral area integration of the EDTA complexes. *J. Solution Chem.* **2004**, *33*, 301–312.

[58] Möbius, K.; Savitsky, A. *High-field EPR Spectroscopy on Proteins and Their Model Systems: Characterization of Transient Paramagnetic States*; Royal Society of Chemistry, 2009.

[59] Bayley, P.; Ahlström, P.; Martin, S. R.; Forsen, S. The kinetics of calcium-binding to calmodulin - quin-2 and ans stopped-flow fluorescence studies. *Biochem. Biophys. Res. Commun.* **1984**, *120*, 185–191.

[60] Martin, S. R.; Linse, S.; Johansson, C.; Bayley, P. M.; Forsen, S. Protein surface charges and calcium binding to individual sites in calbindin D9k: stopped-flow studies. *Biochemistry* **1990**, *29*, 4188–4193.

[61] Martin, S. R.; Andersson Teleman, A.; Bayley, P. M.; Drakenberg, T.; Forsén, S. Kinetics of calcium dissociation from calmodulin and its tryptic fragments. A stopped-flow fluorescence study using Quin 2 reveals a two-domain structure. *Eur. J. Biochem.* **1985**, *151*, 543–550.

[62] Shastry, M. C.; Luck, S. D.; Roder, H. A continuous-flow capillary mixing method to monitor reactions on the microsecond time scale. *Biophys. J.* **1998**, *74*, 2714–21.

[63] Bagchi, S.; Boxer, S. G.; Fayer, M. D. Ribonuclease S Dynamics Measured Using a Nitrile Label with 2D IR Vibrational Echo Spectroscopy. *J. Phys. Chem. B* **2012**, *116*, 4034–4042.

[64] Okuda, M.; Ohta, K.; Tominaga, K. Vibrational Dynamics of Azide-Derivatized Amino Acids Studied by Nonlinear Infrared Spectroscopy. *J. Chem. Phys.* **2015**, *142*, 212418.

[65] van Wilderen, L. J. G. W.; Kern-Michler, D.; Müller-Werkmeister, H. M.; Bredenbeck, J. Vibrational Dynamics and Solvatochromism of the Label SCN in Various Solvents and Hemoglobin by Time Dependent IR and 2D-IR Spectroscopy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 19643–19653.

[66] Shim, S.-H.; Gupta, R.; Ling, Y. L.; Strasfeld, D. B.; Raleigh, D. P.; Zanni, M. T. Two-Dimensional IR Spectroscopy and Isotope Labeling Defines the Pathway of Amyloid

Formation with Residue-Specific Resolution. *Proc. Natl. Acad. Sci.* **2009**, *106*, 6614–6619.

[67] Koziński, M.; Garrett-Roe, S.; Hamm, P. 2D-IR Spectroscopy of the Sulfhydryl Band of Cysteines in the Hydrophobic Core of Proteins. *J. Phys. Chem. B* **2008**, *112*, 7645–7650.

[68] Woys, A. M.; Mukherjee, S. S.; Skoff, D. R.; Moran, S. D.; Zanni, M. T. A Strongly Absorbing Class of Non-Natural Labels for Probing Protein Electrostatics and Solvation with FTIR and 2D IR Spectroscopies. *J. Phys. Chem. B* **2013**, *117*, 5009–5018.

[69] Le Sueur, A. L.; Ramos, S.; Ellefsen, J. D.; Cook, S.; Thielges, M. C. Evaluation of p -(13 C, 15 N-Cyano)phenylalanine as an Extended Time Scale 2D IR Probe of Proteins. *Anal. Chem.* **2017**, *89*, 5254–5260.

[70] Bloem, R.; Koziol, K.; Waldauer, S. A.; Buchli, B.; Walser, R.; Samatanga, B.; Jelesarov, I.; Hamm, P. Ligand Binding Studied by 2D IR Spectroscopy Using the Azidohomoalanine Label. *J. Phys. Chem. B* **2012**, *116*, 13705–13712.

[71] Thielges, M. C.; Axup, J. Y.; Wong, D.; Lee, H. S.; Chung, J. K.; Schultz, P. G.; Fayer, M. D. Two-dimensional IR spectroscopy of protein dynamics using two vibrational labels: A site-specific genetically encoded unnatural amino acid and an active site ligand. *J. Phys. Chem. B* **2011**, *115*, 11294–11304.

[72] Ramos, S.; Mammoser, C. C.; Thibodeau, K. E.; Thielges, M. C. Dynamics underlying hydroxylation selectivity of cytochrome P450cam. *Biophys. J.* **2021**, *120*, 912–923.

[73] Mammoser, C. C.; Ramos, S.; Thielges, M. C. Active Site Hydrogen Bonding Induced in Cytochrome P450cam by Effector Putidaredoxin. *Biochemistry* **2021**, *60*, 1699–1707.

[74] Birdsall, E. R.; Petti, M. K.; Saraswat, V.; Ostrander, J. S.; Arnold, M. S.; Zanni, M. T. Structure Changes of a Membrane Polypeptide under an Applied Voltage Observed with Surface-Enhanced 2D IR Spectroscopy. *J. Phys. Chem. Lett.* **2021**, *12*, 1786–1792.

[75] Ganim, Z.; Tokmakoff, A. Spectral signatures of heterogeneous protein ensembles revealed by MD Simulations of 2DIR spectra. *Biophys. J.* **2006**, *91*, 2636–2646.

[76] La Cour Jansen, T.; Knoester, J. A transferable electrostatic map for solvation effects on amide I vibrations and its application to linear and two-dimensional spectroscopy. *J. Chem. Phys.* **2006**, *124*.

[77] Merchant, K. A.; Thompson, D. E.; Xu, Q.-H.; Williams, R. B.; Loring, R. F.; Fayer, M. D. Myoglobin-CO conformational substate dynamics: 2D vibrational echoes and MD simulations. *Biophys. J.* **2002**, *82*, 3277–3288.

[78] Kratochvil, H. T.; Maj, M.; Matulef, K.; Annen, A. W.; Ostmeyer, J.; Perozo, E.; Roux, B.; Valiyaveetil, F. I.; Zanni, M. T. Probing the Effects of Gating on the Ion

Occupancy of the K+ Channel Selectivity Filter Using Two-Dimensional Infrared Spectroscopy. *J. Am. Chem. Soc.* **2017**, *139*, 8837–8845.

[79] Ganim, Z. 2D IR spectroscopy and computational modeling: application to protein folding and binding. **2010**,

[80] Bagchi, M.; Hochstrasser, 2D-IR Experiments and Simulations of the coupling between amideI and ionizabl eside chains in protein. *NIH* **2012**, *29*, 997–1003.

[81] Baiz, C. R.; McCanne, R.; Nee, M. J.; Kubarych, K. J. Orientational Dynamics of Transient Molecules Measured by Nonequilibrium Two-Dimensional Infrared Spectroscopy. *J. Phys. Chem. A* **2009**, *113*, 8907–8916.

[82] Edington, S. C.; Baiz, C. R. Vibrational Relaxation in EDTA Is Ion-Dependent. *J. Phys. Chem. A* **2018**, *122*, 6585–6592.

[83] Kuroda, D. G.; Vorobyev, D. Y.; Hochstrasser, R. M. Ultrafast relaxation and 2D IR of the aqueous trifluorocarboxylate ion. *J. Chem. Phys.* **2010**, *132*, 044501.

[84] Kuroda, D. G.; Hochstrasser, R. M. Two-dimensional Infrared Spectral Signature and Hydration of the Oxalate Dianion. *J. Chem. Phys.* **2011**, *135*, 1–12.

[85] Ganim, Z.; Tokmakoff, A.; Vaziri, A. Vibrational Excitons in Ionophores: Experimental Probes for Quantum Coherence-assisted Ion Transport and Selectivity in Ion Channels. *New J. Phys.* **2011**, *13*, 113030.

[86] Baiz, C. R.; Peppert, M.; Tokmakoff, A. Amide I Two-Dimensional Infrared Spectroscopy: Methods for Visualizing the Vibrational Structure of Large Proteins. *J. Phys. Chem. A* **2013**, *117*, 5955–5961.

[87] Cunha, A. V.; Bondarenko, A. S.; Jansen, T. L. Assessing Spectral Simulation Protocols for the Amide i Band of Proteins. *J. Chem. Theory Comput.* **2016**, *12*, 3982–3992.

[88] Edington, S. C.; Flanagan, J. C.; Baiz, C. R. An Empirical IR Frequency Map for Ester C=O Stretching Vibrations. **2016**,

[89] Oh, K.-I.; Baiz, C. R. Empirical S=O Stretch Vibrational Frequency Map. *J. Chem. Phys.* **2019**, *151*, 234107.

[90] Baiz, C. R. *et al.* Vibrational Spectroscopic Map, Vibrational Spectroscopy, and Intermolecular Interaction. *Chem. Rev.* **2020**, *120*, 7152–7218.

[91] Bondarenko, A. S.; Jansen, T. L. Application of two-dimensional infrared spectroscopy to benchmark models for the amide I band of proteins. *J. Chem. Phys.* **2015**, *142*.

[92] Barth, A.; Zscherp, C. What Vibrations Tell Us About Proteins. *Q. Rev. Biophys.* **2002**, *35*, 369–430.

[93] Edington, S. C.; Liu, S.; Baiz, C. R. *Infrared spectroscopy probes ion binding geometries.*; 2021; pp 157–191.

[94] Edington, S. C.; Gonzalez, A.; Middendorf, T. R.; Halling, D. B.; Aldrich, R. W.; Baiz, C. R. Coordination to Lanthanide Ions Distorts Binding Site Conformation in Calmodulin. *Proc. Natl. Acad. Sci.* **2018**, *115*, E3126–E3134.

[95] Edington, S. C.; Halling, D. B.; Bennett, S. M.; Middendorf, T. R.; Aldrich, R. W.; Baiz, C. R. Non-Additive Effects of Binding Site Mutations in Calmodulin. *Biochemistry* **2019**, *58*, 2730–2739.

[96] Greaves, T. L.; Weerawardena, A.; Fong, C.; Krodkiewska, I.; Drummond, C. J. Protic Ionic Liquids: Solvents with Tunable Phase Behavior and Physicochemical Properties. *J. Phys. Chem. B* **2006**, *110*, 22479–22487.

[97] Greaves, T. L.; Drummond, C. J. Protic Ionic Liquids: Properties and Applications. *Chem. Rev.* **2008**, *108*, 206–237.

[98] Greaves, T. L.; Weerawardena, A.; Krodkiewska, I.; Drummond, C. J. Protic Ionic Liquids: Physicochemical Properties and Behavior as Amphiphile Self-Assembly Solvents. *J. Phys. Chem. B* **2008**, *112*, 896–905.

[99] Greetham, G. M.; Donaldson, P. M.; Nation, C.; Sazanovich, I. V.; Clark, I. P.; Shaw, D. J.; Parker, A. W.; Towrie, M. A 100 kHz Time-Resolved Multiple-Probe Femtosecond to Second Infrared Absorption Spectrometer. *Appl. Spectrosc.* **2016**, *70*, 645–653.

[100] Brigouleix, C.; Anouti, M.; Jacquemin, J.; Caillon-Caravanier, M.; Galiano, H.; Lemordant, D. Physicochemical Characterization of Morpholinium Cation Based Protic Ionic Liquids Used As Electrolytes. *J. Phys. Chem. B* **2010**, *114*, 1757–1766.

[101] Greaves, T. L.; Drummond, C. J. Protic Ionic Liquids: Evolving Structure–Property Relationships and Expanding Applications. *Chem. Rev.* **2015**, *115*, 11379–11448.

[102] Reid, J. E. S. J.; Bernardes, C. E. S.; Agapito, F.; Martins, F.; Shimizu, S.; Minas da Piedade, M. E.; Walker, A. J. Structure–Property Relationships in Protic Ionic Liquids: A Study of Solvent–Solvent and Solvent–Solute Interactions. *Phys. Chem. Chem. Phys.* **2017**, *19*, 28133–28138.

[103] Stoimenovski, J.; Izgorodina, E. I.; MacFarlane, D. R. Ionicity and Proton Transfer in Protic Ionic Liquids. *Phys. Chem. Chem. Phys.* **2010**, *12*, 10341.

[104] Greaves, T. L.; Kennedy, D. F.; Mudie, S. T.; Drummond, C. J. Diversity Observed in the Nanostructure of Protic Ionic Liquids. *J. Phys. Chem. B* **2010**, *114*, 10022–10031.

[105] Fumino, K.; Wulf, A.; Ludwig, R. Hydrogen Bonding in Protic Ionic Liquids : Reminiscent of Water. *Angew. Chem. Int. Ed.* **2009**, *48*, 3184–3186.

[106] Johnson, C. A.; Parker, A. W.; Donaldson, P. M.; Garrett-Roe, S. An ultrafast vibrational study of dynamical heterogeneity in the protic ionic liquid ethyl-ammonium nitrate. I. Room temperature dynamics. *J. Chem. Phys.* **2021**, *154*, 134502.

[107] Agmon, N. Elementary Steps in Excited-State Proton Transfer †. **2005**, 13–35.

[108] Mohammed, O. F.; Mohammed, O. F.; Dreyer, J.; Magnes, B.-z.; Pines, E. Solvent-Dependent Photoacidity State of Pyranine Monitored by Transient Mid-Infrared Spectroscopy Solvent-Dependent Photoacidity State of Pyranine Monitored by Transient Mid-Infrared Spectroscopy. **2005**,

[109] Ando, K.; Hynes, J. T. *Ionization of Acids in Water.*; 1994; pp 143–153.

[110] Ando, K.; Hynes, J. T. Molecular Mechanism of HCl Acid Ionization in Water: Ab Initio Potential Energy Surfaces and Monte Carlo Simulations. *J. Phys. Chem. B* **1997**, *101*, 10464–10478.

[111] Ando, K.; Hynes, J. T. Molecular Mechanism of HF Acid Ionization in Water: An Electronic Structure-Monte Carlo Study. *J. Phys. Chem. A* **1999**, *103*, 10398–10408.

[112] Kiefer, P. M.; Hynes, J. T. Nonlinear Free Energy Relations for Adiabatic Proton Transfer Reactions in a Polar Environment. I. Fixed Proton Donor-Acceptor Separation. *J. Phys. Chem. A* **2002**, *106*, 1834–1849.

[113] Siwick, B. J.; Bakker, H. J. On the Role of Water in Intermolecular Proton-Transfer Reactions. *J. Am. Chem. Soc.* **2007**, *129*, 13412–13420.

[114] Elsaesser, T.; Becker, H. *Ultrafast Hydrogen Bonding Dynamics and Proton Transfer Processes in the Condensed Phase.*, 3rd ed.; Springer Netherlands, 2013.

[115] Carpenter, W. B.; Fournier, J. A.; Lewis, N. H. C.; Tokmakoff, A. Picosecond Proton Transfer Kinetics in Water Revealed with Ultrafast IR Spectroscopy. *J. Phys. Chem. B* **2018**, *122*, 2792–2802.

[116] Kirchner, B. Eigen or Zundel ion: News From Calculated and Experimental Photoelectron Spectroscopy. *Chemphyschem* **2007**, *8*, 41–43.

[117] Proton, B.; In, T.; Solution, A. Chapter 6 Bimolecular Proton Transfer in Aqueous Solution. **2005**, *301*, 9593–9610.

[118] Mohammed, O. F.; Pines, D.; Dreyer, J.; Pines, E.; Nibbering, E. T. J. Sequential Proton Transfer Through Water Bridges in Acid-Base Reactions. *Science* **2005**, *310*, 83–86.

[119] Mohammed, O. F.; Pines, D.; Pines, E.; Nibbering, E. T. J. Aqueous Bimolecular Proton Transfer in Acid-Base Neutralization. *Chem. Phys.* **2007**, *341*, 240–257.

[120] Greetham, G.; Donaldson, P.; Nation, C.; Sazanovich, I.; Clark, I.; Shaw, D.; Parker, A.; Towrie, M. Time-Resolved Multiple-Probe Spectroscopy on ULTRA-LIFEtime. *CLF Ann. Rep,* **2015**,

[121] Mukamel, S. *Principles of Nonlinear Optical Spectroscopy*; Oxford University Press: New York, 1995.

[122] Hamm, P. Principles of Nonlinear Optical Spectroscopy : A Practical Approach or : Mukamel for Dummies. *Univ. Zurich* **2005**, *41*, 77.

[123] Hamm, P.; Hochstrasser, R. M. In *Ultrafast Infrared Raman Spectrosc.*; Fayer, M. D., Ed.; Marcel Dekker, Inc: New York, 2001; Vol. 26; pp 273–347.

[124] Nitzan, A. *Chemical Dynamics in Condensed Phases*, 1st ed.; Oxford University Press: Oxford, 2006.

[125] Krimm, S.; Abe, Y. Intermolecular Interaction Effects in the Amide I Vibrations of Polypeptides. *Proc. Natl. Acad. Sci.* **1972**, *69*, 2788–2792.

[126] Torii, H.; Tasumi, M. Model calculations on the amide-I infrared bands of globular proteins. *J. Chem. Phys.* **1992**, *96*, 3379.

[127] Moran, A.; Mukamel, S. The Origin of Vibrational Mode Couplings in Various Secondary Structural Motifs of Polypeptides. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 506–510.

[128] Mills, I.; Robiette, A. On the Relationship of Normal Modes to Local Modes in Molecular Vibrations. *Mol. Phys.* **1985**, *56*, 743–765.

[129] Lehmann, K. K. On the relation of Child and Lawton's harmonically coupled anharmonic–oscillator model and Darling–Dennison coupling. *J. Chem. Phys.* **1983**, *79*, 1098–1098.

[130] Jensen, F. *Introduction to Computational Chemistry*, 3rd ed.; Wiley, 2017.

[131] Sholl, D.; Steckel, J. A. *Density Functional Theory: A Practical Introduction*, 2nd ed.; Wiley, 2011.

[132] Hamm, P.; Lim, M.; Hochstrasser, R. M. Vibrational Energy Relaxation of the Cyanide Ion in Water. *J. Chem. Phys.* **1997**, *107*, 10523–10531.

[133] Hamm, P.; Kaindl, R. A.; Stenger, J. Noise Suppression in Femtosecond Mid-Infrared Light Sources. *Opt. Lett.* **2000**, *25*, 1798–1800.

[134] Ren, Z.; Ivanova, A. S.; Couchot-Vore, D.; Garrett-Roe, S. Ultrafast Structure and Dynamics in Ionic Liquids: 2D-IR Spectroscopy Probes the Molecular Origin of Viscosity. *J. Phys. Chem. Lett.* **2014**, *5*, 1541–1546.

[135] Yumoto, F.; Nara, M.; Kagi, H.; Iwasaki, W.; Ojima, T.; Nishita, K.; Nagata, K.; Tanokura, M. Coordination structures of Ca 2+ and Mg 2+ in Akazara scallop troponin C in solution. *Eur. J. Biochem.* **2001**, *268*, 6284–6290.

[136] Nara, M.; Morii, H.; Tanokura, M. Coordination to Divalent Cations by Calcium-binding Proteins Studied by FTIR Spectroscopy. *Biochim. Biophys. Acta - Biomembr.* **2013**, *1828*, 2319–2327.

[137] Christov, C. *Metal-Containing Enzymes*; Advances in Protein Chemistry and Structural Biology; Elsevier Science, 2014.

[138] Dudev, T.; Lim, C. Metal Binding Affinity and Selectivity in Metalloproteins: Insights from Computational Studies. *Annu. Rev. Biophys.* **2008**, *37*, 97–116.

[139] Dudev, T.; Lim, C. Competition among Metal Ions for Protein Binding Sites: Determinants of Metal Ion Selectivity in Proteins. *Chem. Rev.* **2014**, *114*, 538–556.

[140] Dudev, T.; Grauffel, C.; Lim, C. How Pb 2+ Binds and Modulates Properties of Ca 2+ -Signaling Proteins. *Inorg. Chem.* **2018**, *57*, 14798–14809.

[141] Dudev, T.; Ilieva, S.; Doudeva, L. How an electric field can modulate the metal ion selectivity of protein binding sites: insights from DFT/PCM calculations. *Phys. Chem. Chem. Phys.* **2018**, *20*, 24633–24640.

[142] Sawyer, D. T.; Paulsen, P. J. Properties and Infrared Spectra of Ethylenediaminetetraacetic Acid Complexes. I. Alkaline Earth Chelates. *J. Am. Chem. Soc.* **1958**, *80*, 1597–1600.

[143] Sawyer, T.; Tackett, E. Properties and Infrared Spectra of Ethylenediaminetetraacetic Acid Complexes. IV. Structure of Several Metal Chelates in Solution. *J.Phys.Chem* **1960**, *2*, 2390–2394.

[144] Sawyer, D. T.; Tackett, J. E. Properties and Infrared Spectra of Ethylenediaminetetraacetic Acid Complexes. V. Bonding and Structure of Several Metal Chelates in Solution. *J. Am. Chem. Soc.* **1963**, *85*, 2390–2394.

[145] Nakamoto, K.; Morimoto, Y.; Martell, A. E. Infrared Spectra of Aqueous Solutions. III. Ethylenediaminetetraacetic Acid, N-Hydroxyethylethylenediaminetriacetic Acid and Diethylenetriaminepentaacetic Acid. *J. Am. Chem. Soc.* **1963**, *85*, 309–313.

[146] Esteban, M.; Serrano, R.; Vilchez, F. Synthesis and vibrational study of some polydentate ligands. *Spectrochim. Acta Part A Mol. Spectrosc.* **1987**, *43*, 1039–1043.

[147] Mizuguchi, M.; Nara, M.; Kawano, K.; Nitta, K. FT-IR Study of the Calcium-Binding to Bovine α-Lactalbumin. Relationships between the Type of Coordination and Characteristics of the Bands due to the Asp COO- Groups in the Ca2+-Binding Site. *FEBS Lett.* **1997**, *417*, 153–156.

[148] Lanigan, K. C.; Pidsosny, K. Reflectance FTIR Spectroscopic Analysis of Metal Complexation to EDTA and EDDS. *Vib. Spectrosc.* **2007**, *45*, 2–9.

[149] Edington, S. C.; Baiz, C. R. Vibrational Relaxation in EDTA Is Ion-Dependent. *J. Phys. Chem. A* **2018**, *122*, 6585–6592.

[150] Lim, M.; Hamm, P.; Hochstrasser, R. M. Protein Fluctuations are Sensed by Stimulated Infrared Echoes of the Vibrations of Carbon Monoxide and Azide Probes. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 15315–15320.

[151] Drenth, J. *Principles of Protein X-ray Crystallography*; Springer Advanced Texts in Chemistry; Springer New York, 2002.

[152] Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G.; Skelton, N. J.; Rance, M. *Protein NMR Spectroscopy*, 2nd ed.; Elsevier, 2007.

[153] Markwick, P. R. L.; Malliavin, T.; Nilges, M. Structural Biology by NMR: Structure, Dynamics, and Interactions. *PLoS Comput. Biol.* **2008**, *4*, e1000168.

[154] Shi, Y. A Glimpse of Structural Biology through X-Ray Crystallography. *Cell* **2014**, *159*, 995–1014.

[155] Henzler-Wildman, K.; Kern, D. Dynamic personalities of proteins. *Nature* **2007**, *450*, 964–972.

[156] Ghosh, A.; Tucker, M. J.; Hochstrasser, R. M. Identification of Arginine Residues in Peptides by 2D-IR Echo Spectroscopy. *J. Phys. Chem. A* **2011**, *115*, 9731–9738.

[157] Reppert, M.; Tokmakoff, A. Computational Amide I 2D IR Spectroscopy as a Probe of Protein Structure and Dynamics. *Annu. Rev. Phys. Chem.* **2016**, *67*, 359–386.

[158] Deacon, G. B.; Phillips, R. J. Relationships between the Carbon-Oxygen Stretching Frequencies of Carboxylate Complexes and the Typs of Carboxylate Coordination. *Coord. Chem. Rev.* **1980**, *33*, 227–250.

[159] Mizuguchi, M.; Nara, M.; Ke, Y.; Kawano, K.; Hiraoki, T.; Nitta, K. Fourier-Transform Infrared Spectroscopic Studies on the Coordination of the Side-chain COO--groups to Ca2+ in Equine Lysozyme. *Eur. J. Biochem.* **1997**, *250*, 72–76.

[160] Mizuguchi, M.; Fujisawa, R.; Nara, M.; Nitta, K.; Kawano, K. Fourier-Transform Infrared Spectroscopic Study of Ca2+-binding to Osteocalcin. *Calcif. Tissue Int.* **2001**, *69*, 337–342.

[161] Deacon, G.; Huber, F.; Phillips, R. Diagnosis of the nature of carboxylate coordination from the direction of shifts of carbon–oxygen stretching frequencies. *Inorganica Chim. Acta* **1985**, *104*, 41–45.

[162] Jackson, M.; Haris, P. I.; Chapman, D.; Hark, P. I. Fourier Transform Infrared Spectroscopic Studies of Ca(2+)-Binding Proteins. *Biochemistry* **1991**, *30*, 9681–6.

[163] Nara, M.; Tanokura, M. Infrared Spectroscopic Study of the Metal-coordination Structures of Calcium-binding Proteins. *Biochem. Biophys. Res. Commun.* **2008**, *369*, 225–239.

[164] Erwin, N.; Patra, S.; Winter, R. Probing Conformational and Functional Substates of Calmodulin by High Pressure FTIR Spectroscopy: Influence of Ca 2+ Binding and the Hypervariable Region of K-Ras4B. *Phys. Chem. Chem. Phys.* **2016**, *18*, 30020–30028.

[165] Stevenson, P.; Götz, C.; Baiz, C. R.; Akerboom, J.; Tokmakoff, A.; Vaziri, A. Visualizing KcsA Conformational Changes upon Ion Binding by Infrared Spectroscopy and Atomistic Modeling. *J. Phys. Chem. B* **2015**, *119*, 5824–5831.

[166] Strong, S. E.; Hestand, N. J.; Kananenka, A. A.; Zanni, M. T.; Skinner, J. L. IR Spectroscopy Can Reveal the Mechanism of K+ Transport in Ion Channels. *Biophys. J.* **2020**, *118*, 254–261.

[167] Kuroda, D. G.; Hochstrasser, R. M. Dynamic structures of aqueous oxalate and the effects of counterions seen by 2D IR. *Phys. Chem. Chem. Phys.* **2012**, *14*, 6219.

[168] Stefan, I. C.; Mandler, D.; Scherson, D. A. In Situ FTIR-ATR Studies of Functionalized Self-Assembled Bilayer Interactions with Metal Ions in Aqueous Solutions. *Langmuir* **2002**, *18*, 6976–6980.

[169] Donaldson, P. M.; Hamm, P. Gold Nanoparticle Capping Layers: Structure, Dynamics, and Surface Enhancement Measured Using 2D-IR Spectroscopy. *Angew. Chemie Int. Ed.* **2013**, *52*, 634–638.

[170] Flaschka, H. A. *EDTA Titrations: An Introduction to Theory and Practice*; Elsevier Science, 2013.

[171] Helbing, J.; Hamm, P. Compact Implementation of Fourier Transform Two-Dimensional IR Spectroscopy Without Phase Ambiguity. *J. Opt. Soc. Am. B* **2011**, *28*, 171.

[172] Zhao, Y.; Truhlar, D. G. Applications and Validations of the Minnesota Density Functionals. *Chem. Phys. Lett.* **2011**, *502*, 1–13.

[173] Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

[174] Chai, J.-D.; Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **2008**, *128*, 84106.

[175] Shao, Y. *et al.* Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184–215.

[176] Dasgupta, S.; Herbert, J. M. Standard grids for high-precision integration of modern density functionals: SG-2 and SG-3. *J. Comput. Chem.* **2017**, *38*, 869–882.

[177] Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396.

[178] Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminform.* **2012**, *4*, 17.

[179] Sutton, C. C. R.; da Silva, G.; Franks, G. V. Modeling the IR Spectra of Aqueous Metal Carboxylate Complexes: Correlation between Bonding Geometry and Stretching Mode Wavenumber Shifts. *Chem. - A Eur. J.* **2015**, *21*, 6801–6805.

[180] Sutton, C. C.; Franks, G. V.; da Silva, G. Modeling the antisymmetric and symmetric stretching vibrational modes of aqueous carboxylate anions. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2015**, *134*, 535–542.

[181] Katsyuba, S. A.; Zvereva, E. E.; Burganov, T. I. Is There a Simple Way to Reliable Simulations of Infrared Spectra of Organic Compounds? *J. Phys. Chem. A* **2013**, *117*, 6664–6670.

[182] Katsyuba, S. A.; Spicher, S.; Gerasimova, T. P.; Grimme, S. Fast and Accurate Quantum Chemical Modeling of Infrared Spectra of Condensed-Phase Systems. *J. Phys. Chem. B* **2020**, *124*, 6664–6670.

[183] Farrokhpour, H.; Jouypazadeh, H. Description of adenine and cytosine on Au(111) nano surface using different DFT functionals (PW91PW91, wB97XD, M06-2X, M06-L and CAM-B3LYP) in the framework of ONIOM scheme: Non-periodic calculations. *Chem. Phys.* **2017**, *488-489*, 1–10.

[184] Semidalas, E.; Chrissanthopoulos, A. Computational study of structural, vibrational and electronic properties of the highly symmetric molecules M4S6 (M = P, As, Sb, Bi). *Comput. Theor. Chem.* **2019**, *1149*, 41–48.

[185] Horn, P. R.; Mao, Y.; Head-Gordon, M. Probing non-covalent interactions with a second generation energy decomposition analysis using absolutely localized molecular orbitals. *Phys. Chem. Chem. Phys.* **2016**, *18*, 23067–23079.

[186] Horn, P. R.; Mao, Y.; Head-Gordon, M. Defining the contributions of permanent electrostatics, Pauli repulsion, and dispersion in density functional theory calculations of intermolecular interaction energies. *J. Chem. Phys.* **2016**, *144*, 114107.

[187] Mao, Y.; Loipersberger, M.; Kron, K. J.; Derrick, J. S.; Chang, C. J.; Sharada, S. M.; Head-Gordon, M. Consistent inclusion of continuum solvation in energy decomposition

analysis: theory and application to molecular CO2 reduction catalysts. *Chem. Sci.* **2021**, 1398–1414.

[188] Hohenstein, E. G.; Sherrill, C. D. Density fitting and Cholesky decomposition approximations in symmetry-adapted perturbation theory: Implementation and application to probe the nature of π-π interactions in linear acenes. *J. Chem. Phys.* **2010**, *132*, 184111.

[189] Hohenstein, E. G.; Parrish, R. M.; Sherrill, C. D.; Turney, J. M.; Schaefer, H. F. Large-scale symmetry-adapted perturbation theory computations via density fitting and Laplace transformation techniques: Investigating the fundamental forces of DNA-intercalator interactions. *J. Chem. Phys.* **2011**, *135*, 174107.

[190] Smith, D. G. A. *et al.* PSI4 1.4: Open-source software for high-throughput quantum chemistry. *J. Chem. Phys.* **2020**, *152*, 184108.

[191] Shannon, R. D. Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. *Acta Crystallogr. Sect. A* **1976**, *32*, 751–767.

[192] Helgaker, T.; Gauss, J.; Jørgensen, P.; Olsen, J. The prediction of molecular equilibrium structures by the standard electronic wave functions. *J. Chem. Phys.* **1997**, *106*, 6430–6440.

[193] Nakashige, T. G.; Bowman, S. E. J.; Zygiel, E. M.; Drennan, C. L.; Nolan, E. M. Biophysical Examination of the Calcium-Modulated Nickel-Binding Properties of Human Calprotectin Reveals Conformational Change in the EF-Hand Domains and His 3 Asp Site. *Biochem.* **2018**, *57*, 4155.

[194] Halling, D. B.; Liebeskind, B. J.; Hall, A. W.; Aldrich, R. W. Conserved properties of individual Ca 2+ -binding sites in calmodulin. *Proc. Natl. Acad. Sci.* **2016**, *113*, E1216–E1225.

[195] O'boyle, N. M.; Tenderholt, A. L.; Langner, K. M. cclib: A library for package-independent computational chemistry algorithms. *J. Comput. Chem.* **2008**, *29*, 839–845.

[196] Berquist, E.; Hutchison, G.; Langner, K. M.; O'Boyle, N. M.; Tenderholt, A. L.; Upadhyay, S. Release of cclib version 1.7. 2021.

[197] Armand, M.; Endres, F.; MacFarlane, D. R.; Ohno, H.; Scrosati, B. Ionic-Liquid Materials for the Electrochemical Challenges of the Future. *Nat. Mater.* **2009**, *8*, 621–629.

[198] Hayes, R.; Imberti, S.; Warr, G. G.; Atkin, R. The Nature of Hydrogen Bonding in Protic Ionic Liquids. *Angew. Chemie Int. Ed.* **2013**, *52*, 4623–4627.

[199] Ohmine, I.; Tanaka, H. Fluctuation, relaxations, and hydration in liquid water. Hydrogen-bond rearrangement dynamics. *Chem. Rev.* **1993**, *93*, 2545–2566.

[200] Luzar, A.; Chandler, D. Hydrogen-bond kinetics in liquid water. *Nature* **1996**, *379*, 55–57.

[201] Xu, H.; Stern, H. A.; Berne, B. J. Can Water Polarizability Be Ignored in Hydrogen Bond Kinetics? *J. Phys. Chem. B* **2002**, *106*, 2054–2060.

[202] Tarek, M.; Tobias, D. J. Role of Protein-Water Hydrogen Bond Dynamics in the Protein Dynamical Transition. *Phys. Rev. Lett.* **2002**, *88*, 4.

[203] Omta, A. W.; Kropman, M. F.; Woutersen, S.; Bakker, H. J. Negligible effect of ions on the hydrogen-bond structure in liquid water. *Science* **2003**, *301*, 347–349.

[204] Sterpone, F.; Stirnemann, G.; Hynes, J. T.; Laage, D. Water Hydrogen-Bond Dynamics around Amino Acids: The Key Role of Hydrophilic Hydrogen-Bond Acceptor Groups. *J. Phys. Chem. B* **2010**, *114*, 2083–2089.

[205] Yasuda, T.; Watanabe, M. Protic ionic liquids: Fuel cell applications. *MRS Bull.* **2013**, *38*, 560–566.

[206] Penna, T. C.; Faria, L. F. O.; Matos, J. R.; Ribeiro, M. C. C. Pressure and temperature effects on intermolecular vibrational dynamics of ionic liquids. *J. Chem. Phys.* **2013**, *138*, 104503.

[207] Paschoal, V. H.; Faria, L. F. O.; Ribeiro, M. C. C. Vibrational Spectroscopy of Ionic Liquids. *Chem. Rev.* **2017**, *117*, 7053–7112.

[208] UMEBAYASHI, Y.; CHUNG, W.-L.; MITSUGI, T.; FUKUDA, S.; TAKEUCHI, M.; FUJII, K.; TAKAMUKU, T.; KANZAKI, R.; ISHIGURO, S.-i. Liquid Structure and the Ion-Ion Interactions of Ethylammonium Nitrate Ionic Liquid Studied by Large Angle X-Ray Scattering and Molecular Dynamics Simulations. *J. Comput. Chem. Japan* **2008**, *7*, 125–134.

[209] Gontrani, L.; Bodo, E.; Triolo, A.; Leonelli, F.; D'Angelo, P.; Migliorati, V.; Caminiti, R. The Interpretation of Diffraction Patterns of Two Prototypical Protic Ionic Liquids: a Challenging Task for Classical Molecular Dynamics Simulations. *J. Phys. Chem. B* **2012**, *116*, 13024–13032.

[210] Atkin, R.; Warr, G. G. The Smallest Amphiphiles: Nanostructure in Protic Room-Temperature Ionic Liquids with Short Alkyl Groups. *J. Phys. Chem. B* **2008**, *112*, 4164–4166.

[211] Hayes, R.; Imberti, S.; Warr, G. G.; Atkin, R. How Water Dissolves in Protic Ionic Liquids ** Angewandte. **2012**, 1–5.

[212] Hayes, R.; Imberti, S.; Warr, G. G.; Atkin, R. Pronounced sponge-like nanostructure in propylammonium nitrate. *Phys. Chem. Chem. Phys.* **2011**, *13*, 13544–13551.

[213] Hayes, R.; Imberti, S.; Warr, G. G.; Atkin, R. Amphiphilicity determines nanostructure in protic ionic liquids. *Phys. Chem. Chem. Phys.* **2011**, *13*, 3237–3247.

[214] Sonnleitner, T.; Nikitina, V.; Nazet, A.; Buchner, R. Do H-bonds explain strong ion aggregation in ethylammonium nitrate + acetonitrile mixtures? *Phys. Chem. Chem. Phys.* **2013**, *15*, 18445.

[215] Zahn, S.; Thar, J.; Kirchner, B. Structure and dynamics of the protic ionic liquid monomethylammonium nitrate ([CH3NH3][NO3]) from ab initio molecular dynamics simulations. *J. Chem. Phys.* **2010**, *132*, 124506.

[216] Mohammed, O. F.; Pines, D.; Nibbering, E. T. J.; Pines, E. Base-Induced Solvent Switches in Acid – Base Reactions **. **2007**, 1458–1461.

[217] Rini, M.; Magnes, B.-Z.; Pines, E.; Nibbering, E. T. J. Real-Time Observation of Bimodal Proton Transfer in Acid-Base Pairs in Water. *Science* **2003**, *301*, 349–352.

[218] Agmon, N. The Grotthuss mechanism. *Chem. Phys. Lett.* **1995**, *244*, 456–462.

[219] Ueki, T.; Watanabe, M. Macromolecules in Ionic Liquids: Progress, Challenges, and Opportunities. *Macromol. (Washington, DC, U. S.)* **2008**, *41*, 3739–3749.

[220] Nibbering, E. T. J.; Fidder, H.; Pines, E. ULTRAFAST CHEMISTRY: Using Time-Resolved Vibrational Spectroscopy for Interrogation of Structural Dynamics. *Annu. Rev. Phys. Chem.* **2005**, *56*, 337–367.

[221] Spry, D. B.; Goun, A.; Fayer, M. D. Deprotonation Dynamics and Stokes Shift of Pyranine (HPTS). *J. Phys. Chem. A* **2007**, *111*, 230–237.

[222] Spry, D. B.; Goun, A.; Bell, C. B.; Fayer, M. D. Identification and properties of the La1 and Lb1 states of pyranine. *J. Chem. Phys.* **2006**, *125*, 144514.

[223] Greetham, G. M.; Sole, D.; Clark, I. P.; Parker, A. W.; Pollard, M. R.; Towrie, M. Time-resolved multiple probe spectroscopy. *Rev. Sci. Instrum.* **2012**, *83*, 103107.

[224] Sandu, A.; Sander, R. Technical Note: Simulating Chemical Systems in Fortran90 and Matlab with the Kinetic PreProcessor KPP-2.1. *Atmos. Chem. Phys. Discuss.* **2005**, *5*, 8689–8714.

[225] Damian, V.; Sandu, A.; Damian, M.; Potra, F.; Carmichael, G. R. The Kinetic PreProcessor KPP-A Software Environment for Solving Chemical Kinetics. *Comput. Chem. Eng.* **2002**, *26*, 1567–1579.

[226] Cohen, B.; Leiderman, P.; Huppert, D. Unusual Temperature Dependence of Proton Transfer. 2. Excited-State Proton Transfer from Photoacids to Water. *J. Phys. Chem. A* **2002**, *106*, 11115–11122.

[227] Codescu, M.-A.; Weiß, M.; Brehm, M.; Kornilov, O.; Sebastiani, D.; Nibbering, E. T. J. Switching between Proton Vacancy and Excess Proton Transfer Pathways in the Reaction between 7-Hydroxyquinoline and Formate. *J. Phys. Chem. A* **2021**, *125*, 1845–1859.

[228] Hayes, R.; Warr, G. G.; Atkin, R. Structure and Nanostructure in Ionic Liquids. *Chem. Rev.* **2015**, *115*, 6357–6426.

[229] Bruehl, M.; Pan, D.; Ferrer-Vinent, I. J. Demystifying the Chemistry Literature: Building Information Literacy in First-Year Chemistry Students through Student-Centered Learning and Experiment Design. *J. Chem. Educ.* **2015**, *92*, 52–57.

[230] Kim, T. Building Student Proficiency with Scientific Literature Using the Zotero Reference Manager Platform. *Biochem. Mol. Biol. Educ.* **2011**, *39*, 412–415.

[231] Rosenstein, I. J. A Literature Exercise Using SciFinder Scholar for the Sophomore-Level Organic Chemistry Course. *J. Chem. Educ.* **2005**, *82*, 652.

[232] Gawalt, E. S.; Adams, B. A Chemical Information Literacy Program for First-Year Students. *J. Chem. Educ.* **2011**, *88*, 402–407.

[233] Pence, L. E.; Pence, H. E. Accessing and Managing Scientific Literature: Using RSS in the Classroom. *J. Chem. Educ.* **2008**, *85*, 1449.

[234] Hike, N.; Hughes-Phelan, S. J. Using the Science Writing Heuristic to Support NGSS-Aligned Instruction. *J. Chem. Educ.* **2020**, *97*, 358–367.

[235] Bennett, N. S.; Taubman, B. F. Reading Journal Articles for Comprehension Using Key Sentences: An Exercise for the Novice Research Student. *J. Chem. Educ.* **2013**, *90*, 741–744.

[236] Slade, D. J.; Miller, J. S. A Project Provides an Opportunity: Multiple Drafts of an Introduction Require Students To Engage Deeply with the Literature. *J. Chem. Educ.* **2017**, *94*, 1458–1463.

[237] Forest, K.; Rayne, S. Incorporating Primary Literature Summary Projects into a First-Year Chemistry Curriculum. *J. Chem. Educ.* **2009**, *86*, 592.

[238] Tilstra, L. Using Journal Articles to Teach Writing Skills for Laboratory Reports in General Chemistry. *J. Chem. Educ.* **2001**, *78*, 762.

[239] Brownell, S. E.; Price, J. V.; Steinman, L. A Writing-Intensive Course Improves Biology Undergraduates' Perception and Confidence of their Abilities to Read Scientific Literature and Communicate Science. *Adv. Physiol. Educ.* **2013**, *37*, 70–79.

[240] Sivey, J. D.; Lee, C. M. Using Popular Magazine Articles To Teach the Art of Writing for Nontechnical Audiences. *J. Chem. Educ.* **2008**, *85*, 55.

[241] Hoskins, S. G.; Lopatto, D.; Stevens, L. M. The C.R.E.A.T.E. Approach to Primary Literature Shifts Undergraduates' Self-Assessed Ability to Read and Analyze Journal Articles, Attitudes about Science, and Epistemological Beliefs. *CBE—Life Sci. Educ.* **2011**, *10*, 368–378.

[242] Spector, T. I. Managing Undergraduates for Success in the Research Laboratory: Easing the Transition from Routine Class Work to Independent Research. *J. Chem. Educ.* **1993**, *70*, 146.

[243] Drake, B. D.; Acosta, G. M.; Smith, R. L. An Effective Technique for Reading Research Articles - The Japanese KENSHU Method. *J. Chem. Educ.* **1997**, *74*, 186.

[244] Spiegelberg, B. D. A Focused Assignment Encouraging Deep Reading in Undergraduate Biochemistry. *Biochem. Mol. Biol. Educ.* **2014**, *42*, 1–5.

[245] Murray, T. A. Teaching Students to Read the Primary Literature Using Pogil Activities. *Biochem. Mol. Biol. Educ.* **2014**, *42*, 165–173.

[246] Roecker, L. Introducing Students to the Scientific Literature. *J. Chem. Educ.* **2007**, *84*, 1380.

[247] Kovarik, M. L. Use of Primary Literature in the Undergraduate Analytical Class. *Anal. Bioanal. Chem.* **2016**, *408*, 3045–3049.

[248] Hiolski, What's Ice Cream, and Why Do We Scream for It? *Chem. Eng. News* **2018**, 96(31).

[249] Cogné, C.; Laurent, P.; Andrieu, J.; Ferrand, J. Experimental Data and Modelling of Ice Cream Freezing. *Chem. Eng. Res. Des.* **2003**, *81*, 1129–1135.

[250] Fang, C.; Wang, C.; Xiong, Y. L.; Pomper, K. W. Extraction and Characterization of Polyphenol Oxidase in Pawpaw (*asiminatriloba*) Fruit. *J. Food Biochem.* **2007**, *31*, 603–620.

[251] Yilsay, T. Ö.; Yilmaz, L.; Bayizit, A. A. The Effect of Using a Whey Protein Fat Replacer on Textural and Sensory Characteristics of Low-Fat Vanilla Ice Cream. *Eur. Food Res. Technol.* **2006**, *222*, 171–175.

[252] Zulim Botega, D. C.; Marangoni, A. G.; Smith, A. K.; Goff, H. D. Development of Formulations and Processes to Incorporate Wax Oleogels in Ice Cream. *J. Food Sci.* **2013**, *78*, C1845–C1851.

[253] Madhusoodanan, J. Laser-Heated Nanoparticles Bust Eye "Floaters". *Chem. Eng. News* **2019**, July 23 (special to C&EN).

[254] Sauvage, F.; Fraire, J. C.; Remaut, K.; Sebag, J.; Peynshaert, K.; Harrington, M.; Van de Velde, F. J.; Xiong, R.; Tassignon, M.-J.; Brans, T.; Braeckmans, K.; De

Smedt, S. C. Photoablation of Human Vitreous Opacities by Light-Induced Vapor Nanobubbles. *ACS Nano* **2019**, *13*, 8401–8416.

[255] Skirtach, A. G.; Antipov, A. A.; Shchukin, D. G.; Sukhorukov, G. B. Remote Activation of Capsules Containing Ag Nanoparticles and IR Dye by Laser Light. *Langmuir* **2004**, *20*, 6988–6992.

[256] Baffou, G.; Polleux, J.; Rigneault, H.; Monneret, S. Super-Heating and Micro-Bubble Generation around Plasmonic Nanoparticles under cw Illumination. *J. Phys. Chem. C* **2014**, *118*, 4890–4898.

[257] Pitsillides, C. M.; Joe, E. K.; Wei, X.; Anderson, R. R.; Lin, C. P. Selective Cell Targeting with Light-Absorbing Microparticles and Nanoparticles. *Biophys. J.* **2003**, *84*, 4023–4032.

[258] Kurachi, Y.; North, A. Ion Channels: Their Structure, Function and Control - An Overview. *J. Physiol.* **2004**, *554*, 245–247.

[259] Garrido, E.; Lázaro, J.; Jaumot, M.; Agell, N.; Rubio-Martinez, J. Modeling and Subtleties of K-Ras and Calmodulin Interaction. *PLOS Comput. Biol.* **2018**, *14*, e1006552.

[260] Cho, H. Regulation of Adenosine-Activated GIRK Channels by Gq-Coupled Receptors in Mouse Atrial Myocytes. *Korean J. Physiol. Pharmacol.* **2010**, *14*, 145.

[261] Griffith, T.; Tsaneva-Atanasova, K.; Mellor, J. R. Control of Ca2+ Influx and Calmodulin Activation by SK-Channels in Dendritic Spines. *PLOS Comput. Biol.* **2016**, *12*, e1004949.

[262] Gammon, K. Amino Acids Could Help Determine Sex Of Fingerprint Originators. *Chem. Eng. News* **2015**,

[263] Ifa, D. R.; Manicke, N. E.; Dill, A. L.; Cooks, R. G. Latent Fingerprint Chemical Imaging by Mass Spectrometry. *Science* **2008**, *321*, 805–805.

[264] Agudelo, J.; Halámková, L.; Brunelle, E.; Rodrigues, R.; Huynh, C.; Halámek, J. Ages at a Crime Scene: Simultaneous Estimation of the Time since Deposition and Age of Its Originator. *Anal. Chem.* **2016**, *88*, 6479–6484.

[265] Day, J. S.; Edwards, H. G.; Dobrowski, S. A.; Voice, A. M. The Detection of Drugs of Abuse in Fingerprints using Raman Spectroscopy II: Cyanoacrylate-Fumed Fingerprints. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2004**, *60*, 1725–1730.

[266] Chen, T.; Schultz, Z. D.; Levin, I. W. Infrared Spectroscopic Imaging of Latent Fingerprints and Associated Forensic Evidence. *Analyst* **2009**, *134*, 1902.

[267] Hasib-ur Rahman, M.; Siaj, M.; Larachi, F. Ionic liquids for CO2 capture—Development and progress. *Chem. Eng. Process. Process Intensif.* **2010**, *49*, 313–322.

[268] Ren, Z.; Brinzer, T.; Dutta, S.; Garrett-Roe, S. Thiocyanate as a Local Probe of Ultrafast Structure and Dynamics in Imidazolium-Based Ionic Liquids: Water Induced Heterogeneity and Cation Induced Ion-Pairing. *J. Phys. Chem. B* **2015**, *119*, 4699–4712.

[269] Cadena, C.; Anthony, J. L.; Shah, J. K.; Morrow, T. I.; Brennecke, J. F.; Maginn, E. J. Why Is CO 2 So Soluble in Imidazolium-Based Ionic Liquids? *J. Am. Chem. Soc.* **2004**, *126*, 5300–5308.

[270] Wells, A. S.; Coombe, V. T. On the Freshwater Ecotoxicity and Biodegradation Properties of Some Common Ionic Liquids. *Org. Process Res. Dev.* **2006**, *10*, 794–798.

[271] Ribeiro, C. P.; Freeman, B. D. Sorption, Dilation, and Partial Molar Volumes of Carbon Dioxide and Ethane in Cross-Linked Poly(ethylene oxide). *Macromolecules* **2008**, *41*, 9458–9468.

[272] Hoffmann, F.; Ekimova, M.; Bekçioğlu-Neff, G.; Nibbering, E. T. J.; Sebastiani, D. Combined Experimental and Theoretical Study of the Transient IR Spectroscopy of 7-Hydroxyquinoline in the First Electronically Excited Singlet State. *J. Phys. Chem. A* **2016**, *120*, 9378–9389.

[273] Ekimova, M.; Hoffmann, F.; Bekçioğlu-Neff, G.; Rafferty, A.; Kornilov, O.; Nibbering, E. T. J.; Sebastiani, D. Ultrafast Proton Transport between a Hydroxy Acid and a Nitrogen Base along Solvent Bridges Governed by the Hydroxide/Methoxide Transfer Mechanism. *J. Am. Chem. Soc.* **2019**, *141*, 14581–14592.

[274] Hayes, R.; Imberti, S.; Warr, G. G.; Atkin, R. How Water Dissolves in Protic Ionic Liquids. *Angew. Chemie - Int. Ed.* **2012**, *51*, 7468–7471.

[275] Hayes, R.; Imberti, S.; Warr, G. G.; Atkin, R. Effect of cation alkyl chain length and anion type on protic ionic liquid nanostructure. *J. Phys. Chem. C* **2014**, *118*, 13998–14008.

[276] Sabnis, R. *Handb. Fluoresc. Dye. Probes*; John Wiley & Sons, Inc: Hoboken, NJ, USA, 2015; Chapter 90, pp 253–262.